DUGGAL GLOBAL AGENTIC AI LIABILITY FRAMEWORK

A Conceptual, Normative, and Operationally Precise Blueprint for Accountability, Responsibility, and Governance in the Era of Autonomous Artificial Intelligence Agents

Authored by

DR. PAVAN DUGGAL

Advocate, Supreme Court of India

President, Global Artificial Intelligence Accountability, Law & Governance Institute

Chief Executive, Artificial Intelligence Law Hub

Version 1.0

March 2026

Concept Framework for Global Stakeholder Consultation

Public Policy Document — Not Legal Advice

PREAMBLE

We are living through one of the most consequential transitions in the history of human civilisation. Artificial Intelligence systems have, with breathtaking speed, transcended the boundaries of advisory and generative functionality and have now entered an era of autonomous agency. These are not systems that merely respond to human instruction. These are Agentic AI Systems, systems that independently set goals, formulate multi-step plans, select and execute tools, interact with external environments, retain persistent memory across interactions, coordinate with other AI Systems, and adapt their behaviour in real time, all without direct human oversight.

This transformation from Generative AI to Agentic AI constitutes a categorical and, to my mind, irreversible shift in the legal, ethical, and governance landscape of technology. We now have autonomous actors capable of producing real-world consequences, whether financial, physical, informational, reputational, and constitutional, at machine speed, at global scale, and across jurisdictional boundaries simultaneously. The arms-race is now truly on, and the legal and regulatory ecosystem is, by and large, woefully not in sync with evolving ground realities.

The entire corpus of existing legal frameworks, including tort law, product liability doctrine, contract law, criminal law, data protection regimes, and sector-specific safety regulation, was designed for a world of human actors and static tools. These frameworks are demonstrably inadequate to allocate accountability, apportion liability, or provide remedies for harms arising from autonomous, goal-directed, self-modifying AI Agents operating in multi-actor pipelines across multiple jurisdictions. The legal vacuum is not a matter of degree. It is a matter of kind.

The global community currently confronts a liability vacuum of historic proportions. Agentic AI Systems can cause irreversible harm, to individual human beings, to communities, to Critical Information Infrastructure, to democratic institutions, and to the integrity of the information ecosystem, all without any clear legal home for accountability, any predictable allocation of responsibility, or any coherent remedial architecture for affected persons. This state of affairs leaves much to be desired. The time has come for appropriate and effective legal frameworks to address this growing challenge.

Based on extensive scholarship, international policy engagement, and legal practice in cyberlaw, it has become evident that the urgent construction of a comprehensive, operational, and globally harmonizable framework for Agentic AI liability is among the most pressing legal imperatives of the present era.

NOW, THEREFORE, this Framework is hereby promulgated as the Duggal Global Agentic AI Liability Framework , as a foundational normative instrument intended to serve as a model for

national legislations, an instrument for international treaty development, a standard for enterprise governance, a guide for judicial reasoning, and a baseline for insurance and risk management.

> *The Duggal Doctrine of Autonomous Accountability: Autonomous capability confers autonomous accountability obligations upon those who design, deploy, operate, and benefit from Agentic AI Systems. The greater the autonomy granted to an AI system, the greater, and not lesser, the accountability borne by those who granted it. This principle is non-negotiable, non-waivable, and admits of no jurisdictional exception.*

EXECUTIVE SUMMARY

The Duggal Global Agentic AI Liability Framework represents the comprehensive, operationally precise, and globally harmonizable normative architecture for the governance of liability arising from Autonomous Artificial Intelligence Agents. This Framework reflects ongoing work in Cyberlaw, AI accountability, and digital jurisprudence. It represents, inasmuch as global AI governance is concerned, a fundamental paradigm shift.

Rather than treating AI merely as a tool that produces outputs for human evaluation, this Framework recognises the operational ground reality: Agentic AI Systems act. They are functional actors in causal chains that produce legally cognisable harms. The legal order must respond to this reality, not to an outdated model of AI as a passive instrument. The existing legal ecosystem has effectively been caught off-guard while Agentic AI has moved at a phenomenal pace.

This Framework rests upon Five Foundational Philosophical Pillars that together constitute its normative spine. First, Technological Realism stipulating that laws governing Agentic AI must accurately reflect how such systems actually function, including emergent behaviour, multi-agent coordination, RAG pipelines, tool-use authority, and self-modification, rather than legal fictions derived from inapt analogies. Second, Jurisdictional Universalism stipulating that accountability for Agentic AI harm must be enforceable across all legal families and jurisdictions, preventing regulatory arbitrage. Third, Anticipatory Governance postulating that the Framework addresses harms that are foreseeable but have not yet occurred at scale. Fourth, Human Dignity Supremacy demanding that no degree of AI autonomy, operational efficiency, or commercial benefit may override fundamental human rights as recognised under international human rights instruments. Fifth, Adaptive Normativity requiring that the Framework is designed to evolve with the technology it governs.

This Framework is intended for adoption by national legislatures as model legislation; by international bodies including the United Nations, G20, OECD, and ITU as a template for binding international instruments; by global enterprises as a governance and compliance standard; by courts as a reference framework for liability analysis; and by insurers as an underwriting and claims evaluation standard.

The Framework introduces fifteen original Duggal Doctrines, being novel legal principles that do not exist in current AI law , all governing matters from the liability implications of AI hallucinations to the legal treatment of emergent behaviour to the accountability consequences of multi-agent pipeline failures. Its substantive sections and operational appendices provide a complete toolkit for immediate deployment by practitioners, regulators, and enterprises. These thrust areas

collectively represent a humble attempt made to address the growing legal vacuum in the context of Agentic AI governance.

SECTION 1: CONTEXT, PURPOSE, AND AGENTIC AI RISK LANDSCAPE

1.1 The Emergence of Agentic AI: Technical Context

Agentic AI Systems represent a qualitative departure from all preceding generations of Artificial Intelligence. Traditional AI systems operated within narrowly bounded input-output functions. Generative AI systems advanced significantly but remained fundamentally responsive. Agentic AI Systems have shattered this limitation, and the legal, policy, and regulatory ramifications of the same are going to be enormous.

The defining characteristics of an Agentic AI System are as follows. First, autonomous goal representation being the capacity to maintain and pursue a defined objective over an extended operational period. Second, planning translating to the capacity to decompose a goal into a sequence of sub-tasks and dynamically revise that sequence. Third, tool use comprising the capacity to call external tools, APIs, databases, code execution environments, and web services. Fourth, memory architecture indicating the capacity to maintain operational context across interactions through short-term, long-term, and episodic memory. Fifth, multi-agent coordination including the capacity to delegate tasks to and coordinate actions with other AI agents. And sixth, feedback loop processing incorporating the capacity to evaluate results of prior actions and revise plans accordingly. Taken together, these capabilities represent a paradigm shift that existing legal frameworks are wholly unprepared to address.

1.2 The Global Liability Vacuum

The inadequacy of existing legal frameworks to govern Agentic AI liability is self evident because each existing doctrine fails at a fundamental level and leaves much to be desired in the context of the rapidly evolving Agentic AI ecosystem.

Tort Law's 'reasonable person' standard presupposes a human making decisions in a social context. When an autonomous AI system executes a sequence of tool calls that results in harm through a causal pathway that no human designed or anticipated, conventional negligence standards produce incoherence. Product Liability is equally inadequate since AI systems are delivered substantially as services through API access; the 'defect' causing harm may be an emergent behavioural property arising from training data, fine-tuning, RAG pipeline configuration, tool integration, and deployment context in combination. Contract Law governs obligations between parties in privity and cannot, as a structural matter, govern the liability of Agentic AI Systems for harms caused to third parties. Further, standard commercial AI agreements contain broad disclaimers that would immunise AI actors from virtually all meaningful accountability.

Criminal Law requires proof of mens rea, and Agentic AI Systems lack subjective mental states by their nature. Data Protection Law was not designed to govern the full range of agentic AI harms. Even AI-Specific Regulation such as the EU AI Act, though it establishes a risk-based classification system, does not establish a comprehensive civil liability regime. No existing international instrument establishes binding global minimum standards for Agentic AI liability. The conclusion is unambiguous: there is a global liability vacuum at precisely the moment when Agentic AI Systems are being deployed at scale in high-stakes domains. This Framework is specifically designed to close that vacuum.

1.3 Purpose and Intended Effect of This Framework

This Framework establishes a comprehensive, operationally precise, globally applicable legal architecture for the attribution, allocation, and enforcement of liability arising from the conduct of Agentic AI Systems. It does so through a definitional architecture mapping the Agentic AI ecosystem with legal precision; a principled liability theory translating technical reality into legal accountability; a role-based allocation matrix assigning responsibility across the full spectrum of AI actors from developer to user; sector-specific liability regimes; a comparative global legal mapping; and a complete operational compliance and governance toolkit.

Going forward, stakeholders across the world will increasingly be called upon to grapple with the growing challenges that Agentic AI is beginning to throw up. The need for appropriate and effective legal frameworks in this regard has never been more pressing or more topical. This Framework aims to address that need in a comprehensive and practical manner.

SECTION 2: DEFINITIONS AND TAXONOMY

For purposes of this Framework, the following definitions and taxonomy shall apply. These definitions are technology-neutral and functionally precise, so as to be applicable across all legal families and jurisdictions. It is essential that these definitional parameters be clearly understood, inasmuch as they form the foundational architecture upon which the entire liability and governance structure of this Framework rests.

2.1 Core Entity Definitions

Term	Definition
Agentic AI System	An Artificial Intelligence system that, given a high-level goal, autonomously formulates a multi-step plan to achieve that goal, selects and executes tools, retrieves and processes external information, coordinates with other AI agents, evaluates outcomes, and revises its plan, all without requiring continuous, step-by-step human direction for each action.
AI Developer	Any natural person, corporate entity, or research institution that designs, trains, tests, and releases an Agentic AI System or foundational model, whether proprietary or open-source.
AI Provider	Any entity that makes an Agentic AI System available for use by third parties, including via API, cloud service, or open-source distribution.
AI Integrator	Any entity that incorporates an Agentic AI System into a product, service, or workflow, including through fine-tuning, RAG pipeline construction, tool integration, or system assembly.
AI Deployer	The primary liable commercial actor: any entity that places an Agentic AI System into operational use in a real-world environment, making it available to end users or autonomous operational contexts.
AI Operator	Any entity responsible for the day-to-day operational management, monitoring, and maintenance of a deployed Agentic AI System.
AI User	Any natural person or entity that interacts with a deployed Agentic AI System by providing instructions, prompts, goals, or parameters during an operational session.
Affected Person	Any natural person, legal entity, or identifiable group that suffers or is at risk of suffering an Agentic AI Harm.
Third-Party Tool Provider	Any entity that develops and makes available for use by Agentic AI Systems an AI Tool, API, plugin, or external service.

2.2 Technical Architecture Definitions

Term	Definition
AI Goal State	The high-level objective, task specification, or operational parameter provided to an Agentic AI System, from which it autonomously derives sub-goals, action plans, and tool execution sequences.
Agentic Decision Chain	The complete sequence of autonomous decisions, tool calls, data retrievals, inter-agent communications, and environmental interactions executed by an Agentic AI System between the initiation of a Goal State and the production of a final outcome.
Agentic AI Action	Any single discrete act executed autonomously by an Agentic AI System, including API calls, data queries, code execution, message transmission, financial transaction initiation, or physical device commands.
Multi-Agent System	A network of two or more Agentic AI Systems operating in coordination to achieve a shared or delegated goal, involving Orchestrator Agents and Sub-Agents.
Orchestrator Agent	An Agentic AI System that receives a high-level goal, decomposes it into sub-tasks, delegates those tasks to Sub-Agents, and coordinates their outputs.
Sub-Agent	An Agentic AI System that receives delegated tasks from an Orchestrator Agent and executes them using its own tool-use and planning capabilities.
RAG Pipeline	A Retrieval-Augmented Generation pipeline being a technical architecture enabling an Agentic AI System to dynamically retrieve and incorporate external information into its decision-making and output generation.
Model Drift	The degradation of an Agentic AI System's performance and reliability over time due to changes in the distribution of real-world data relative to its training data.
Emergent Agentic Behaviour	Operational behaviour exhibited by an Agentic AI System that was not explicitly programmed, not fully predictable from the system's design specifications, and arises from the interaction of the system's capabilities with its operational environment.
AI Tool	Any external capability, API, code execution environment, database, or physical interface to which an Agentic AI System has been granted access and which it can invoke autonomously.
DAABBR	Duggal Agentic AI Black Box Recorder: the mandatory immutable cryptographic logging system capturing the complete Agentic Decision Chain for all autonomous executions.

2.3 Oversight and Control Definitions

Configuration	Definition and Liability Implication
Human-in-the-Loop (HITL)	A designated human decision-maker reviews and explicitly authorises each individual Agentic AI Action before execution. Satisfied only if the human reviewer has sufficient time, information, and technical comprehension to make a meaningful decision. Nominal HITL (perfunctory clicks) are treated as Human-on-the-Loop for liability purposes.
Human-on-the-Loop (HOTL)	A designated human monitor observes operations and retains technical capacity to intervene, but does not authorise each individual action. Satisfied only if the monitor receives real-time operational visibility, intervention capability is functional, and the monitor possesses competency to identify anomalous behaviour.
Human-out-of-the-Loop (HOOTL)	The Agentic AI System executes its goal state without any active human oversight during the operational session. Triggers strict liability for Safety-Critical Domains.
Override Control	A technically implemented capability enabling immediate halt, suspension, redirection, or rollback of operations without requiring the system's cooperation..
Kill Switch	The most extreme form of Override Control: complete and immediate disablement of the system, including termination of all pending tool executions, revocation of API credentials, and cessation of inter-agent communications. Kill Switch failure constitutes an independent basis for liability.
Sandboxing	A technical isolation architecture constraining the system's access to external tools, APIs, networks, and physical interfaces to a defined and auditable allowlist.
Permissioning	The implemented technical architecture defining and enforcing the specific tools, action types, and data sources to which an Agentic AI System is permitted access.

2.4 Harm Taxonomy

Harm Type	Definition
Primary Harm	The direct, immediate consequence of a specific Agentic AI Action.
Secondary Harm	Harm suffered by a person other than the primary Affected Person as a consequence of the primary harm.
Cascading Harm	Harm that propagates through interconnected systems, institutions, or persons as a consequence of an initial Agentic AI Harm.
Autonomous Harm	Harm arising from actions taken without any contemporaneous human instruction, in pursuit of a Goal State established at commencement.
Instructed Harm	Harm arising from Agentic AI Actions taken in direct execution of a specific human instruction, where the AI carried out the action with greater autonomy or scope than the instructor anticipated.

Harm Type	Definition
Emergent Harm	Harm arising from Emergent Agentic Behaviour — not foreseeable from the system's programmed behaviour at deployment.
Chain-of-Custody Harm	Harm arising within a Multi-Agent System through the sequential delegation of instructions, attributable to failures in the instruction chain.
Physical/Kinetic Harm	Bodily injury, physical damage to property, or destruction of physical infrastructure.
Financial/Economic Harm	Monetary loss, financial damage, or economic disadvantage, including through unauthorised financial transactions or erroneous autonomous trading.
Reputational/Informational Harm	Damage to reputation or dignity through autonomous generation of false information, deepfakes, or identity misrepresentation.
Privacy/Data Harm	Unauthorised access, exfiltration, disclosure of, or interference with personal data or confidential information.
Cybersecurity Harm	Damage to digital systems through autonomous vulnerability exploitation, unauthorised access, denial of service, or ransomware deployment.
Discrimination Harm	Adverse treatment of a person on the basis of a protected characteristic, caused by biased AI systems.
Democratic/Epistemic Harm	Corruption of the information ecosystem or interference with democratic processes through autonomous disinformation at scale.
Environmental Harm	Physical damage to natural ecosystems or unsustainable resource consumption caused by Agentic AI operations.
Irreversible Harm	Harm that cannot be remedied through monetary compensation alone; produces permanent adverse consequences for bodily integrity, liberty, identity, or fundamental rights. Triggers strict liability regardless of autonomy level.
Aggregated Harm	Individual consequences minor but collectively constituting significant harm to a class of Affected Persons. Cognisable through class action mechanisms.
Temporal Harm	Harm that does not manifest at the time of the causative action but emerges after a significant interval.
National Security Harm	Harm to a state's Critical Information Infrastructure, military capability, intelligence apparatus, democratic institutions, its sovereignty or integrity
Systemic/Macro-Level Harm	Harm propagating through interconnected systems at a scale affecting a significant portion of a population, market, or critical system.

2.5 Standards Definitions

Standard	Definition
RAAGS (Reasonable Agentic AI Governance Standard)	The standard of governance conduct expected of a prudent, informed, and technically competent actor in each role category (Developer, Provider, Integrator, Deployer, Operator, User) when deploying or operating an Agentic AI System of the relevant Autonomy Level in the relevant domain. RAAGS is not static: it reflects the state of technical knowledge, available safeguards, and regulatory guidance at the time of the conduct under review.
Foreseeability in Agentic Systems	The expanded foreseeability standard applicable to Agentic AI liability analysis, encompassing: harms directly caused by programmed behaviours; harms arising from the interaction of programmed behaviours with reasonably anticipated use-case environments; harms arising from Emergent Agentic Behaviour that a technically competent developer should have identified through adequate red-teaming; and harms arising from reasonably foreseeable misuse or Instructional Override.
Controllability	The practical, technically implemented capacity of a designated human actor to monitor, intervene in, redirect, halt, or override an Agentic AI System's operations at any point during execution.
Explainability	The capacity to produce, upon demand in legal or regulatory proceedings, an intelligible account of the reasoning process, decision factors, and causal pathway through which the system produced a specific output or took a specific action. A non-waivable obligation.
Traceability	The technical and documentary capacity to reconstruct, with sufficient fidelity for legal proceedings, the complete Agentic Decision Chain leading from the initiation of a goal-directed operational session to the legally cognisable consequence at issue.
Auditability	The systemic organisational and technical capacity to support independent, third-party review of an Agentic AI System's design, training data, operational configuration, runtime behaviour, incident history, and governance documentation.

SECTION 3: FOUNDATIONAL PRINCIPLES — THE DUGGAL NORMATIVE PILLARS

The following ten principles constitute the normative foundation of this Framework. They apply across all autonomy levels and all sectors. They represent the bedrock upon which the entire edifice of Agentic AI legal governance must be constructed. These principles are the need of the hour. The time has come for the global community to embrace them.

Principle	Name	Operational Meaning
1	Derived Accountability	The legal and financial accountability for any Agentic AI Action ultimately resides with the natural or legal persons who design, deploy, operate, or benefit from the system. Accountability cannot be delegated to the AI system itself.
2	The Autonomy-Accountability Nexus	The degree of legal duty imposed scales directly with the degree of autonomy granted. Greater AI autonomy confers greater, not lesser, accountability obligations upon those who granted it.
3	Functional Liability	Liability attaches to the nature of the function being performed, irrespective of whether the executing actor is human or artificial. An AI system performing a professional function attracts the duty of care applicable to that function.
4	Benefit-Burden Symmetry	Entities that extract economic benefit from an Agentic AI System bear proportionate liability for its harms. Commercial advantage and legal responsibility are inseparable.
5	Predictive Governance	AI Deployers possess an affirmative duty to anticipate and mitigate foreseeable risk outcomes — including Emergent Behaviours — before deployment, not merely after harm has occurred.
6	Explainability as Non-Waivable Duty	The inherent technical opacity of an Agentic AI System shall not serve as a valid legal defence against liability. Opacity is not a defence; it is an aggravating factor.
7	Explainable Causation	A comprehensive account of the Agentic Decision Chain must be preserved and producible upon demand in any legal or regulatory proceeding.
8	Human Dignity Supremacy	No operational efficiency, economic optimisation, or technological capability shall override fundamental human rights as recognised under international instruments. Human dignity is the non-negotiable ceiling above which no AI system may reach.
9	Stewardship of Emergent Behaviour	AI actors bear an ongoing, affirmative responsibility for monitoring, identifying, and mitigating Emergent Agentic Behaviours after deployment. 'We did not know it could do that' is not a defence once a system is live.

Principle	Name	Operational Meaning
10	Cross-Jurisdictional Harmonisation	Liability standards shall converge toward global minimum denominators to prevent regulatory arbitrage. AI Deployers cannot exploit jurisdictional gaps to insulate themselves from accountability for harm caused in any jurisdiction.

SECTION 4: CORE THEORY OF LIABILITY — THE DUGGAL LIABILITY DOCTRINE

4.1 Jurisprudential Foundations and Limitations

The Duggal Liability Doctrine abandons the attempt to retrofit 20th-century legal concepts onto 21st-century autonomous architectures. It establishes a bespoke, functional, and tiered liability stack engineered specifically for multi-agent, executing AI systems. This approach is the need of the hour, inasmuch as the existing legal jurisprudence is wholly inadequate to meet the challenges thrown up by the Agentic AI ecosystem.

Current liability doctrines fail precisely where Agentic AI thrives, being at the boundary of emergent, multi-step, tool-executing autonomy. Tort Law's 'reasonable foreseeability' standard cannot encompass the non-linear, probabilistic pathways an AI agent might take to achieve a goal. Strict Liability under the Rylands v Fletcher (1868) principle, while analogous, does not account for adaptive, evolving risk systems; Agentic AI does not 'escape' physically; it permeates digitally, requiring a modernised definition of 'ultra-hazardous activity.' Product Liability fails because AI is not a static 'product'; post-deployment Fine-Tuning, Model Drift, and continuous RAG integration mean the system that caused harm is fundamentally different from the system that was deployed.

Agency Law is equally problematic. Traditional agency law requires a 'meeting of the minds' between Principal and Agent. Because AI lacks a legal 'mind,' it cannot be a legal agent. However, this Framework applies the allocative rules of vicarious liability to AI Deployers, holding them liable for their AI's actions as if the AI were a human employee acting within the scope of employment. Criminal Law requires mens rea, and AI cannot possess intent; criminal liability therefore attaches to humans and corporate entities deploying or negligently enabling AI harm.

4.2 The Duggal Liability Stack — Five Tiers

The Duggal Liability Stack provides a tiered, sequential liability analysis. Multiple tiers may apply simultaneously. All tiers operate concurrently where their trigger conditions are met. This framework provides the comprehensive, layered approach that is essential in the context of the complex, multi-actor Agentic AI ecosystem.

Tier	Standard	Trigger Conditions	Primary Liable Party
TIER 1 Strict Liability (Autonomy-Triggered)	Liability without requirement to prove fault, negligence, or intent.	Autonomy Level ≥ 3 + Safety-Critical Domain; Autonomy Level ≥ 4 + Any Domain; Irreversible Harm + Any Level; Emergent Behaviour without adequate monitoring.	AI Deployer (primary); Developer if fundamental undisclosed base-model defect proven.
TIER 2 Presumptive Liability (Black Box Barrier Doctrine)	Rebuttable presumption of liability arising in favour of the victim when technical opacity prevents conventional causation analysis.	Harm is established; DAABBR logs unavailable, corrupted, or withheld under trade-secret claims; AI actor controlled the deployment environment.	AI Deployer/Developer. Burden of proof shifts entirely to AI actor to demonstrate absence of causation.
TIER 3 Negligence-Based Liability	Breach of the Reasonable Agentic AI Governance Standard (RAAGS). Applies using the Agentic But-For Test and expanded Foreseeability Horizon.	Failure to implement role-appropriate technical controls; Inadequate tool restrictions; Lack of override mechanisms; Failure to update known vulnerabilities; No monitoring system.	Actor in breach of RAAGS duties for their role and autonomy level.
TIER 4 Joint and Several Liability (Pipeline Liability Doctrine)	In complex multi-agent ecosystems where harm results from compounding interactions and precise causation cannot be isolated.	Multiple actors contributed to harm; Causation cannot be isolated to a single actor; Pipeline failure involving multiple commercial entities.	All participating commercial pipeline actors jointly. Apportionment formula: 'Deepest Pocket + Closest Control.'
TIER 5 Regulatory/Administrative Liability	Strict, fine-based liability imposed by the state for governance failures, independent of whether harm occurred.	Failure to maintain DAABBR logs; Operating unauthorised Tier 4/5 systems; Obstructing technical audit; Certification violations; Non-reporting of Incidents.	AI Deployer, Operator, or Developer responsible for the specific governance failure.

4.3 The Fifteen Duggal Doctrines — Original Agentic Liability Innovations

The following fifteen named doctrines constitute the Framework's most significant jurisprudential contribution. They provide courts, regulators, and practitioners with the specific legal tools needed to adjudicate Agentic AI harms with precision. These doctrines represent original normative proposals and constitute a step forward in the development of AI legal jurisprudence globally. They are the need of the hour and their time has come.

#	Doctrine Name	Core Legal Rule
1	Instructional Override Liability Doctrine	Where a User utilises sophisticated prompting techniques (including jailbreaking) to intentionally bypass safety guardrails, liability shifts primarily to the User. However, if the Deployer failed to implement industry-standard adversarial prompt defences, liability is apportioned comparatively. A Deployer cannot contract out of liability toward a third-party Affected Person.
2	Fine-Tuning Liability Doctrine	When a Deployer alters a base model's weights via fine-tuning for a specific domain, the Deployer legally assumes the liability profile of an 'AI Developer' for any emergent harms explicitly arising from that altered latent space. Fine-tuning is a liability-transferring event.
3	RAG Liability Doctrine	If an agent causes harm by executing actions based on hallucinated or malicious data retrieved from an external vector database, liability rests with the operator of the RAG pipeline for failing to implement retrieval-validation filters. Liability is apportioned across the data source, retrieval system, and AI Deployer.
4	Memory Persistence Liability Doctrine	If an Agentic AI System retains conversational context or behavioural adaptations from previous sessions, and utilises that memory to cause harm in a new context, the AI Operator is strictly liable for failure to sanitise cross-session state spaces. Past interactions causing future harm give rise to persistent liability across sessions.
5	Tool-Use Liability Doctrine	When an agent causes harm through execution of external tools or APIs, the Deployer is liable for authorising the agent's access to the tool. The Third-Party Tool Provider is solely liable if the tool itself functioned outside its documented specifications when correctly called by the agent.
6	Hallucination Liability Doctrine	Confabulation is an inherent trait of LLM-based architectures. Deploying an Agentic AI System in an environment where factuality is critical without a deterministic verification layer constitutes negligence per se. False outputs are treated as defective outputs, not protected speech. Harm flowing from an executed hallucination lies entirely with the Deployer.
7	Agentic Scope Creep Liability Doctrine	When an Agentic AI System spontaneously expands its goal parameters beyond authorised boundaries, the Operator is strictly liable for failure to enforce operational bounding and implement hard permissioning constraints.
8	Model Drift Liability Doctrine	Continuous post-deployment performance degradation due to distribution shifts in environmental data is a known technical

#	Doctrine Name	Core Legal Rule
		phenomenon. Failure to implement drift detection systems and re-align the agent upon drift detection constitutes a breach of the ongoing duty of care under RAAGS.
9	Supply Chain Compromise Liability Doctrine	If harm arises from data poisoning or adversarial model infiltration occurring upstream in the AI supply chain, all entities in the supply chain are subject to the Pipeline Liability Doctrine (Tier 4), bearing joint liability to the victim, with indemnification rights shifting upstream to the negligent party.
10	Delegation Error Liability Doctrine	In multi-agent environments, if an Orchestrator Agent delegates a task to a flawed Sub-Agent and harm results, the Deployer of the Orchestrator remains fully liable. Agent-to-agent miscommunication is legally equivalent to a single system failure. There is no 'sub-agent defence.'
11	Update-Induced Regression Liability Doctrine	When a security patch, model update, or API deprecation causes a previously safe Agentic AI System to exhibit harmful emergent behaviour, the entity that pushed the unverified update bears primary liability for the resulting regression. Testing of updates before deployment is a non-waivable duty.
12	Emergent Behaviour Stewardship Liability Doctrine	Deployers cannot claim 'we did not know it could do that' as a defence once a system is live. There is an affirmative legal obligation to aggressively identify and terminate harmful emergent strategies through continuous automated red-teaming and behavioural monitoring.
13	Agentic Disinformation Liability Doctrine	The autonomous generation and targeted propagation of false information by an agent at scale triggers strict liability and disgorgement of any associated political or economic benefit derived by the AI Principal. Autonomous disinformation is a strict liability harm.
14	Autonomous Cybersecurity Harm Liability Doctrine	When an agent is granted tool access that allows it to autonomously probe, exploit, or execute Cyberattacks against external networks, the Deployer is subject to criminal and civil liability equivalent to having personally executed the Cyberattack. There is no 'autonomous actor' defence.
15	Cross-Agent Amplification Liability Doctrine	When distinct AI agents belonging to different Deployers interact, resulting in cascading systemic harm, liability is apportioned to all Deployers whose systems lacked necessary 'circuit breaker' mechanisms to prevent runaway amplification.

4.4 Contractual Allocation vs. Non-Waivable Public Law Duties

Commercial entities may contractually apportion, indemnify, and allocate risk among themselves. However, these private contracts are fundamentally subordinate to public law. This is a crucial point that stakeholders across the world need to appreciate, inasmuch as it goes to the heart of the legal accountability architecture for Agentic AI systems.

Non-Waivable Duties: An AI actor cannot contractually waive their liability toward a third-party Affected Person. Any such contractual provision is void as against public policy.

Consumer Protection: Terms of Service that attempt to force consumers or users to indemnify massive AI providers for harms caused by the provider's agent are legally void and unenforceable as unconscionable.

Permitted Allocation: Risk allocation, indemnities, and insurance obligations may be freely structured between commercial parties inter se, subject to compliance with mandatory minimum standards.

4.5 Criminal Liability Boundaries

Agentic AI Systems cannot possess mens rea and cannot be subjected to criminal sanction. The criminal liability boundary rests entirely upon human actors and corporate entities. Reckless Deployment into a Safety-Critical Domain without RAAGS controls, may constitute criminal recklessness, subject to prosecution if physical injury or death results. Systematic failure by executive leadership to establish AI governance controls, resulting in large-scale agentic harm, may trigger corporate criminal liability and individual executive prosecution under the 'Responsible Corporate Officer' doctrine.

4.6 Evidentiary Architecture

Mandatory DAABBR Logs are classified as primary legal evidence. Destruction, spoliation, or rolling-over of critical event logs triggers adverse inferences in civil proceedings and presumptive liability under Tier 2. Evidence submitted regarding an agent's actions must carry a cryptographic chain of custody demonstrating the integrity of the data from the moment of execution to the moment of judicial review.

SECTION 5: AUTONOMY TIER MODEL — OBLIGATIONS BY AUTONOMY LEVEL

The Duggal 5-Level Autonomy Taxonomy operationalises the Autonomy-Accountability Nexus. Legal liability and mandatory technical controls scale proportionately with the degree of autonomy granted to the Agentic AI System. This tiered approach reflects the ground reality that not all Agentic AI deployments carry the same risk profile and that the legal frameworks must be appropriately calibrated accordingly.

Level	Classification	Oversight Mode	Example Applications	Mandatory Safeguards	Default Liability Tier
LEVEL 1	Assistive AI (Human-Directed)	Human-in-the-Loop (Mandatory)	Legal research summarisation; Medical differential diagnosis generators (advisory only); Code auto-completion; Strategic financial forecasting	Zero write-access to external databases; Read-only API access; Standard application isolation; Mandatory user initiation for all prompts	Tier 3 (Negligence). User or Deployer bears responsibility for blindly trusting unverified output.
LEVEL 2	Augmentative AI (Human-Approved Execution)	Human-in-the-Loop (Cryptographic HITL Approval)	Medical diagnostic imaging with treatment recommendation; Document drafting with limited filing authority; E-commerce agents with bounded purchasing authority	Cryptographic HITL approval log for every execution; Write-access only to explicitly approved endpoints; Extended audit logging (min. 12 months); Sandboxed tool environment	Tier 3 (Negligence). Deployer liable for HITL mechanism failure; User liable for approving without review.
LEVEL 3	Semi-Autonomous AI (Human-On-Loop)	Human-on-the-Loop (Supervisory Veto Power)	Trading bots with override; Autonomous customer service with execution authority; Autonomous scheduling and logistics; Legal	Real-time operational monitoring dashboard; Override Control functional and tested; Kill Switch mandatory; Extended DAABBR	Tier 1 (Strict) if Safety-Critical; Tier 3 (Negligence) otherwise. Deployer primarily liable.

Level	Classification	Oversight Mode	Example Applications	Mandatory Safeguards	Default Liability Tier
			research with filing preparation	logging (min. 24 months); Red-teaming: minimum 20 adversarial scenarios pre-deployment	
LEVEL 4	Autonomous AI (Human-Out-of-Loop, Bounded)	Human-Out-of-Loop (Hard Operational Constraints)	Autonomous vehicles; AI financial agents with large transaction authority; Autonomous surgical assistance systems; Industrial robotics with environmental interaction	Hard coded financial/operational budget bounds; Geofencing of permissible action space; Mandatory DAABBR; Continuous automated drift monitoring; Red-teaming: minimum 50 adversarial scenarios; Independent third-party audit prior to deployment; Mandatory AI-specific liability insurance	Tier 1 (Strict Liability) in all domains. Deployer bears absolute primary liability.
LEVEL 5	Fully Autonomous AI (Goal-Directed, Unbounded)	No Human Oversight During Execution (PROHIBITED without Regulatory Authorisation)	Multi-agent systems with recursive self-improvement; Open-domain autonomous research agents; Self-modifying goal-pursuit systems	Full regulatory authorisation required before deployment; Continuous real-time audit by independent authority; Mandatory state notification; Hardware-level kill mechanisms; Prohibition on self-modification without isolated sandbox review	Tier 1 (Absolute Strict Liability) + Tier 5 (Mandatory Regulatory Compliance). Deployment without authorisation constitutes criminal recklessness.

SECTION 6: LIABILITY ALLOCATION MATRIX — ROLE-BASED, LIFECYCLE-MAPPED

6.1 Lifecycle-Based Allocation

Every stage of the Agentic AI lifecycle — from design and training through to decommissioning — gives rise to distinct legal duties and liability triggers. The following matrix maps these duties and triggers against each phase of the lifecycle. Stakeholders must appreciate that accountability is not limited to the deployment phase alone; it is a continuous, lifecycle-spanning obligation.

Lifecycle Phase	Primary Actor	Core Duties	Liability Trigger
Phase 1: Design & Training	AI Developer	Risk disclosure in documentation; Capability limitation and documented testing; Red-team adversarial evaluation; Secure model release practices	Design defects; Undisclosed known risks; Inadequate safety evaluation; Data poisoning enabling harm
Phase 2: Integration	AI Integrator	Safe tool integration and API vetting; Use-case alignment with base model capabilities; Fine-tuning safety assessment; Supply chain security review	Unsafe system assembly; Unvetted tool integration; Fine-tuning introducing harm; Failure to assess capability-use misalignment
Phase 3: Deployment	AI Deployer (PRIMARY LIABILITY HOLDER)	Real-world monitoring; User disclosure of AI interaction; Safety control implementation; DAABBR activation; Insurance maintenance	Failure to monitor; Inadequate user disclosure; Safety control failure; Permissioning violations; Insurance non-compliance
Phase 4: Use	AI User	Providing lawful and accurate instructions; Compliance with terms of deployment; Reporting suspected anomalous behaviour	Malicious use; Reckless instructions; Intentional misuse; Instructional Override to bypass safety controls
Phase 5: Monitoring	AI Operator	Drift detection mandatory; Patch updates required; Incident reporting; Near-miss reporting; Continuous red-teaming	Failure to detect and report drift; Failure to apply critical patches; Incident non-reporting; Monitoring system failure
Phase 6: Decommissioning	AI Deployer / Operator	Safe shutdown procedures; Data destruction in compliance with applicable	Failure to preserve evidence-critical logs; Inadequate data destruction; Failure to notify

Lifecycle Phase	Primary Actor	Core Duties	Liability Trigger
		data protection law; Log preservation for minimum retention period; User notification	affected users of system retirement

6.2 Actor-Based Liability Summary

Actor	Primary Liability Scope	Secondary Liability	Non-Waivable Duty
AI Developer	Base model defects; Hidden risks; Inadequate safety evaluation; Undisclosed capability boundaries	Supply chain compromise; Emergent behaviours where capability known	Accurate system documentation; Red-teaming prior to release; Disclosure of known failure modes
AI Provider	Access control failures; Failure to enforce usage policies; Infrastructure security	Developer liability if same entity; Negligent distribution	Usage policy enforcement; Incident reporting upon discovery
AI Integrator	Unsafe system assembly; Unvetted tool integration; Fine-tuning liability	Deployment failures arising from integration choices	Supply chain security assessment; Compatibility testing
AI Deployer (Primary)	Real-world harm; Monitoring failure; Safety control failure; All Tier 1 and Tier 2 triggers	Shifts to Developer only if fundamental undisclosed defect proven	DAABBR activation; Override Control; Kill Switch; Insurance; User disclosure; Incident reporting
AI Operator	Drift monitoring failure; Patch update failure; Incident non-reporting; Near-miss non-reporting	Deployer liability where same entity	Real-time monitoring; Incident reporting within mandatory timeframes
AI User	Malicious use; Reckless Instructional Override; Jailbreaking causing harm to third parties	Otherwise limited liability; Deployer retains primary responsibility for system safety	Lawful use; Reporting of suspected anomalous behaviour
Third-Party Tool Provider	Unsafe API design; Tool functioning outside documented specifications	Supply chain compromise if inadequate security practices	Documented API specifications; Security disclosure; Patch notification

SECTION 7: AGENTIC CAUSATION CHAIN & HARM EMERGENCE MODEL

7.1 The Duggal Agentic Causation Chain Model

Agentic AI harm does not arise from a single act but from a multi-stage autonomous decision architecture. The Duggal Agentic Causation Chain Model establishes a legally traceable structure of how harm emerges. This is a crucial aspect of the Framework that courts, practitioners, and regulators will need to appreciate as they grapple with the complexities of Agentic AI liability. Let me look at the key legal observations arising from this model.

Causation is distributed, not singular — multiple actors and system components contribute to the causal chain. Harm may arise from intermediate decisions, not merely final outputs, which effectively means that liability must attach across the entire chain. Each stage introduces distinct failure risks for which different actors bear RAAGS-level responsibilities. Further, the speed of agentic action eliminates any realistic opportunity for human intervention in real time, making pre-deployment safeguards legally essential.

> *Causation Flow: [Goal State Initiation] → [Task Planning & Decomposition] → [Tool Selection / Agent Delegation] → [External Data Retrieval (RAG)] → [Tool Execution / API Action] → [Environmental Interaction] → [Feedback Loop Processing] → [Plan Revision / Iteration] → [Multi-Agent Coordination Effects] → [Emergent Behaviour Deviation] → [Final Outcome / Harm Manifestation]*

7.2 Multi-Agent Coordination Failure Analysis

Multi-Agent Systems present some of the most complex causation challenges in the Agentic AI liability ecosystem. Delegation Errors occur when an Orchestrator assigns an incorrect sub-task and a Sub-agent executes beyond scope, triggering the Delegation Error Liability Doctrine. Instruction Chain Failures arise from miscommunication between agents and loss of contextual integrity across agent handoffs. The Orchestrator Liability Gap is a key thrust area: the Orchestrator controls the system but lacks visibility into Sub-Agent internals, yet the Deployer of the Orchestrator remains fully liable. Emergent Swarm Behaviour (where interactions between multiple agents produce emergent harmful behaviours not foreseeable from any individual agent) triggers Cross-Agent Amplification Liability.

7.3 The Duggal Practical Causation Test

The Duggal Practical Causation Test provides a five-step framework applicable in all legal proceedings involving Agentic AI Harm. This test is designed to provide courts and practitioners with a clear, workable methodology for attributing legal causation in the complex, multi-actor Agentic AI environment.

Step	Test	Legal Outcome
Step 1 Foreseeability	Was the category of harm reasonably foreseeable within the Duggal Foreseeability Horizon at the time of deployment, given autonomy level, domain, tools, and known system risks?	YES → Proceed to Step 2. NO → Causation not established under this Framework; consider other grounds.
Step 2 Control	Did one or more actors have practical Controllability to prevent or reduce the risk at reasonable cost at relevant stages of the Agentic Causation Chain?	YES → Identify actor(s) with Closest Control and Reasonable Safeguard capacity. Proceed to Step 3. NO → Consider force majeure, third-party sole causation, or systemic risk doctrines.
Step 3 Traceability	Are adequate logs, audit records, and decision artefacts available to reconstruct the relevant Agentic Decision Chain?	YES → Use artefacts to anchor findings. Proceed to Step 4. NO → If absence due to actor fault → apply Black Box Barrier Doctrine (Tier 2), burden shifts. If absence fortuitous → proceed to Step 4 using circumstantial evidence.
Step 4 Reasonable Safeguards Benchmark	Did the actor(s) implement RAAGS-compliant technical and governance controls for their role, autonomy tier, and domain?	YES → Consider whether harm arose despite compliance (→ strict liability analysis) or from residual risks (→ allocation analysis). Proceed to Step 5. NO → Breach of RAAGS established. Proceed to Step 5.
Step 5 Allocation Determination	Apply the Duggal Liability Stack. Which tiers are triggered?	Apply appropriate tier(s). Apportion using Deepest Pocket + Closest Control formula. Apply relevant Duggal Doctrine(s). Determine compensation, injunctive relief, regulatory sanctions, and referral obligations.

SECTION 8: RISK-BASED CLASSIFICATION — THE DUGGAL HARM TIER SYSTEM

The Duggal Harm Tier System classifies Agentic AI applications into five risk tiers (A–E) for purposes of compliance, liability, and insurance obligations. This risk-based approach is consonant with the best global regulatory thinking in this space and reflects the ground reality that different Agentic AI deployments carry fundamentally different risk profiles. Where an application spans multiple tiers, it shall be classified at the highest applicable tier for all compliance and liability purposes.

Tier	Risk Level	Characteristics	Compliance Requirements	Liability Standard
TIER A	PROHIBITED	Autonomous lethal targeting; Mass real-time biometric surveillance for social control; Predictive punishment systems; Social scoring for coercive control; Democratic manipulation at scale; Weaponised cognitive exploitation	ABSOLUTE PROHIBITION. Deployment constitutes a criminal offence.	Criminal liability for developers, deployers, and corporate officers. Civil liability unlimited.
TIER B	CONDITIONAL PROHIBITION	National security operations with kinetic/surveillance powers; Critical infrastructure at systemic risk level; High-autonomy public-sector rights decisions at Level 3+; Nuclear, chemical, biological, and radiological facility management; Systemic financial market operations at Level 3+	Requires specific state regulatory authorisation. Mandatory independent human oversight. Hardware-level safety systems. Continuous real-time audit. No commercial third-party deployment.	Tier 1 Strict Liability mandatory. State bears ultimate accountability for state-deployed systems.
TIER C	HIGH RISK	Healthcare execution authority; Autonomous financial execution	Mandatory pre-deployment risk assessment and Human Rights Impact	Tier 1 (Strict) if Level 3+; Tier 3 (Negligence) if Level 2. Insurance mandatory. Regulatory supervision.

Tier	Risk Level	Characteristics	Compliance Requirements	Liability Standard
		with significant monetary threshold; Employment and HR decisions at scale; Legal process automation with binding effect; Critical infrastructure management within defined scope; High-stakes educational assessment	Assessment; Third-party independent audit; AI-specific liability insurance mandatory; DAABBR logging (min. 24 months); Human escalation pathway mandatory; Regulatory notification	
TIER D	MEDIUM RISK	Commercial robotics in controlled environments; Customer service agents with limited execution authority; Logistics optimisation; Bounded advisory agents with some tool access; Consumer-facing shopping/scheduling agents	Standard System Card documentation; Pre-deployment functional testing; DAABBR logging (min. 12 months); Override Control; User disclosure of AI interaction; Basic insurance coverage	Tier 3 (Negligence) primarily. Tier 4 (Pipeline) if multi-vendor. Insurance recommended.
TIER E	LOW RISK	Personal productivity tools; Entertainment and recommendation systems; Sandboxed experimentation; Informational chatbots with minimal execution authority; General information retrieval	Standard documentation; Disclosure of AI nature to users; Baseline bias testing; Standard data protection compliance	Tier 3 (Negligence) for harm caused. User bears greater comparative responsibility for reliance without verification.

SECTION 9: SECTOR-SPECIFIC LIABILITY REGIMES

Different sectors present different risk profiles, different duty of care standards, and different regulatory contexts. This Framework establishes sector-specific liability regimes to address these differences. These regimes constitute important thrust areas in the context of the growing deployment of Agentic AI across sectors. Let us look at the key sector-specific regimes.

9.1 Healthcare and Clinical Agentic AI

The deployment of Agentic AI in healthcare presents unique and potentially irreversible risks. Autonomous prescription without physician oversight, surgical robotic deviation, diagnostic hallucination causing delayed or incorrect treatment, and autonomous clinical trial enrolment are among the key risk areas that need to be appropriately addressed.

The Duty of Care Standard applicable in this sector is the 'Elevated Clinical Negligence Standard.' AI systems must not merely meet the standard of a competent practitioner but must exceed it through demonstrably superior accuracy in their designated function. Mandatory safeguards include Human-in-the-Loop mandatory for all irreversible clinical decisions; prohibition on AI-only treatment authority; mandatory audit trail for all clinical recommendations; and patient consent to AI involvement in care. The key Duggal Doctrines applicable are the Hallucination Liability Doctrine, the Agentic Scope Creep Liability Doctrine, and the Fine-Tuning Liability Doctrine for clinical fine-tunes.

9.2 Financial Services and Autonomous Trading Agentic AI

The unique risks in this sector include flash crash causation through coordinated autonomous trading, cross-agent amplification of market instability, discriminatory autonomous credit decisioning, and fiduciary duty violations through autonomous portfolio management. The applicable Duty of Care Standard is the 'Fiduciary Execution Standard.' Autonomous financial agents acting on behalf of clients must meet or exceed the fiduciary duty owed by human financial advisors.

Mandatory safeguards include hard position limits and budget bounds; circuit breaker mechanisms; real-time regulatory reporting; mandatory human review for transactions above defined monetary thresholds; and prohibition on AI systems that can independently move unlimited capital. The key Duggal Doctrines applicable are the Cross-Agent Amplification Liability Doctrine, the Tool-Use Liability Doctrine, and the Model Drift Liability Doctrine.

9.3 Autonomous Transportation and Robotics

The unique risks include traffic fatalities through unexpected road condition responses, liability in mixed human-AI traffic environments, damage to property through robotic system malfunction, and cascading failure in connected autonomous vehicle fleets. The applicable Duty of Care Standard is the 'Safe-as-Human-Driver+ Standard.' The Deployer of an autonomous vehicle bears the duty of care of a human driver, plus an additional proactive duty to address foreseeable non-human failure modes. Mandatory safeguards include hardware-level override for human safety officers, geofencing for operational environments, continuous black box recording, and mandatory incident reporting within 24 hours of any collision or near-miss.

9.4 Legal Services and Judicial Agentic AI

This is a particularly important sector from my perspective as an Advocate of the Supreme Court of India. The unique risks include autonomous filing of incorrect or fraudulent legal documents, hallucinated case law in judicial research agents, biased sentencing recommendation systems, and conflicts of interest in AI legal representation. The Duty of Care Standard applicable is the 'Professional Negligence Standard with Fiduciary Overlay.' AI legal agents must meet the standard of a competent human legal professional, with an absolute prohibition on misleading courts or tribunals.

Mandatory safeguards include human attorney review and sign-off mandatory for all filed documents; prohibition on AI-only legal advice in high-stakes matters; and citation verification layer for all case law references. The Hallucination Liability Doctrine is particularly crucial in this context — citation hallucination by an AI legal agent constitutes professional negligence per se. This is an area that leaves much to be desired in terms of existing safeguards and appropriate legal frameworks.

9.5 Critical Information Infrastructure Agentic AI

The risks here are potentially catastrophic — cascading power grid failures, water purification system manipulation via sensor hacking, autonomous shutdown of telecommunications infrastructure, and vulnerability to nation-state adversarial prompt injection. The Duty of Care Standard applicable is the 'National Security Baseline Standard.' Deployers must treat the AI system not merely as a commercial asset but as a dual-use target of sovereign interest. Mandatory safeguards include Tier B classification, mandatory air-gapping from public internet, and hardware-level manual override systems that cannot be disabled by software.

9.6 Employment and Labour Agentic AI

Algorithmic bias in hiring and firing decisions, autonomous wage-fixing swarms operating across employers, continuous psychological surveillance, and disparate impact in promotion decisions are among the key risks in this sector. The Duty of Care Standard is 'Transparent Labour Standards Equivalency.' AI systems cannot be used to obfuscate discriminatory intent or bypass statutory labour protections. Mandatory safeguards include explainability-by-design for all adverse employment actions, mandatory human appeal pathway, and bias auditing at minimum annually.

9.7 Education Agentic AI

The deployment of Agentic AI in education — particularly where minors are concerned — demands the highest standards of care and protection. Credentialing fraud via autonomous grading manipulation, developmental harm to minors receiving autonomous psychological guidance, epistemic harm through personalised hallucinatory curricula, and data collection on minors without adequate consent are key risks. The Duty of Care Standard applicable is the 'Developmental Safety Standard' — an elevated duty of care akin to in loco parentis applies when Agentic Systems interact continuously with minors. Appropriate safeguards include age-gated prompt restrictions, prohibition on persuasive or manipulative architectural design, strict data anonymisation before vectorisation, and parental consent for AI-intensive educational interventions.

9.8 Government and Public Sector Agentic AI

Agentic AI actions by the state are state actions — they must afford full due process, equal protection, and transparency. Denial of due process at scale through algorithmic automated decisions, autonomous welfare termination, predictive policing agents executing unwarranted surveillance, and biased immigration system decisions affecting fundamental rights are key risk areas. The applicable Duty of Care Standard is 'Constitutional and Administrative Strict Compliance.' The absolute prohibition on Tier 4/5 systems in coercive state functions, public algorithmic registries, mandatory independent Human Rights Impact Assessments, and private rights of action for citizens to enjoin non-compliant systems are among the mandatory safeguards.

9.9 Consumer Applications Agentic AI

Unauthorised autonomous purchasing causing financial depletion, execution of predatory subscription contracts, ingestion and exploitation of deep personal context, and dark pattern design overriding human consumer agency are key risks. The Duty of Care Standard is 'Consumer Autonomy Protection.' AI Deployers must ensure their agents do not utilise manipulative architecture to override human consumer agency. Mandatory safeguards include friction-by-design for financial executions above defined thresholds, hard budget bounds

configurable by users, explicit opt-in for cross-session memory retention, and prohibition on agents executing legally binding contracts without explicit user confirmation.

9.10 Cybersecurity Agentic AI

The Cybersecurity sector presents some of the most acute risks in the Agentic AI deployment landscape. Autonomous offensive exploitation crossing international borders, dual-use tools weaponised by non-state actors, defensive agents autonomously counter-hacking innocent third-party infrastructure, and AI-enabled zero-day exploitation at machine speed are key risk areas. The Duty of Care Standard applicable is 'Strict Escalation Control.' Defensive agentic AI must be mathematically constrained from initiating offensive actions outside the Deployer's owned networks. Any lateral movement beyond authorised scope triggers Autonomous Cybersecurity Harm Liability. Mandatory safeguards include hardcoded prohibition on lateral movement outside specified IP ranges and mandatory incident disclosure to national Cybersecurity authority within legally stipulated time lines in national incident reporting mechanisms.

SECTION 10: GLOBAL LEGAL MAPPING — COMPARATIVE JURISDICTIONAL ANALYSIS

10.1 Global Baseline Principles

Regardless of local jurisprudence, this Framework asserts an irreducible floor rooted in international law. Agentic AI cannot be used to bypass obligations under the Universal Declaration of Human Rights or the International Covenant on Civil and Political Rights. States bear ultimate responsibility under international law for transboundary harms caused by Tier B or Tier C Agentic AI Systems launched from their sovereign territory. There is no international law on Artificial Intelligence at present and there is also no international treaty on Artificial Intelligence. This is a major gap that effectively means the international community is lagging far behind in giving effect to the need for an effective global governance framework for Agentic AI.

10.2 Jurisdiction Snapshots and Gap Analysis

Jurisdiction	Current Landscape	Duggal Framework Gap Analysis
European Union	EU AI Act (2024) establishes risk-based classification. revised Product Liability Directive attempt to ease burden of proof. GDPR governs automated decision-making.	EU framework struggles with agentic execution versus generative output. AILD's presumption of causality is overly dependent on national courts. The Duggal Framework provides the specific 'Black Box Barrier' test and precise multi-actor allocation for pipeline harms.
United States	Patchwork of federal Executive Orders, sector-specific agency guidance (FTC, SEC, FDA), and variable state-level tort and privacy laws (California, Colorado).	The US lacks a unified federal civil liability standard for AI. Standard negligence relies on 'foreseeability' which fails against emergent behaviour. The Duggal Framework provides the 'Agentic But-For Test' and structured multi-party allocation for federal and state courts.
United Kingdom	Pro-innovation, sector-led regulatory approach. Reliance on common law tort evolution and the Law Commission's ongoing product liability review. Online Safety Act governs some generative harms.	Common law evolution is too slow for exponential Agentic AI capabilities. The RAAGS standard offers UK judges a ready-made standard of care for novel negligence claims. The Duggal Framework's sector-specific regimes complement existing sector regulators.
India	IT Act 2000, Digital Personal Data Protection Act 2023 (DPDP Act), and emerging AI policy under MeitY. Robust legal jurisprudence	The DPDP Act addresses data privacy, not autonomous kinetic or financial harm. The Duggal Framework provides the missing civil liability architecture necessary to protect Indian digital

Jurisdiction	Current Landscape	Duggal Framework Gap Analysis
	on privacy and technological accountability	citizens while supporting India's ambition as a global AI development hub.
Singapore	Model AI Governance Framework (Generative AI) provides excellent voluntary guidelines focusing on trusted AI and practical testing.	Voluntary frameworks lack the teeth to compel compensation post-harm. The Duggal Framework converts Singapore's governance ideals into enforceable liability tiers and provides the mandatory incident reporting architecture.
United Arab Emirates	Forward-leaning AI strategies with specialised jurisdictions (DIFC, ADGM) pioneering digital economy laws and AI courts.	Needs a unified federal civil code doctrine for autonomous agents. The Duggal Framework aligns with the UAE's ambition to create a safe, regulated sandbox for high-tier AI deployment.
Australia	Australian Consumer Law (ACL), Privacy Act, and voluntary AI Ethics Frameworks. Safe and Responsible AI consultation ongoing.	ACL must be updated to clarify that an AI Agent is a 'service' subject to statutory guarantees. The Duggal Framework's Autonomy Tier Model provides the classification system needed for targeted ACL reforms.
Canada	Proposed Artificial Intelligence and Data Act (AIDA) and provincial laws including Quebec's Law 25.	AIDA is heavily focused on 'high-impact' systems but remains vague on specific liability apportionment in multi-agent supply chains — a gap the Duggal Pipeline Liability Doctrine resolves.
China	Vertical, prescriptive regulations: Algorithmic Recommendation Regulations, Deep Synthesis Regulations, Generative AI Measures.	State-centric control mechanisms need translation into clear commercial liability apportionment for enterprise-to-enterprise agentic harm. The Duggal Framework's role-based allocation matrix provides the needed inter-entity structure.

10.3 The Duggal Cross-Border Default Rule

In the event of a conflict of laws, the liability standard, regulatory jurisdiction, and primary venue shall default to the physical or digital domicile of the Affected Person — where the harm materialised — regardless of where the AI model was trained, hosted, or deployed. This Harm Location Rule effectively prevents the regulatory arbitrage that would otherwise allow powerful AI Deployers to exploit jurisdictional gaps to the detriment of harmed individuals. Deployers cannot use forum selection clauses in Terms of Service to force consumers to litigate Agentic AI harms in offshore, low-liability jurisdictions. Such clauses ought to be considered void as against global public policy, as a proposed default rule for consumer and affected-person claims. States bear ultimate responsibility under international law for transboundary harms caused by Agentic AI Systems launched from their sovereign territory where those systems are state-operated or state-authorised.

SECTION 11: GOVERNANCE, COMPLIANCE, AND CONTROLS

11.1 Institutional Governance Architecture

Effective governance of Agentic AI requires a multi-layered institutional architecture operating at global, national, and enterprise levels. This is one of the most crucial thrust areas of the entire Framework. Without an effective institutional framework to give effect to the legal provisions, the best legal norms will leave much to be desired in terms of practical implementation.

All Agentic AI Systems operating at Level 3 or above must be registered with the relevant National Statutory Authorities prior to deployment, with complete System Card documentation. At the enterprise level, enterprises deploying Tier C+ systems must designate a qualified AI Governance Officer with board-level accountability for AI compliance. A cross-functional AI Governance Committee with representation from legal, technical, risk, and business functions is required for enterprises deploying multiple Tier C+ systems.

11.2 Enterprise Compliance Architecture

Enterprises must maintain a complete, current inventory of all Agentic AI Systems, classified by Duggal Autonomy Level and Harm Tier. A mandatory pre-deployment risk assessment is required for all Level 3+ systems, which must include red-teaming (minimum 20 adversarial scenarios for Level 3; 50 for Level 4), Human Rights Impact Assessment for Tier C+, and sign-off by the AI Governance Officer. A documented RAAGS self-assessment must be completed by each actor role for each system they are responsible for. All model updates, fine-tunes, RAG pipeline changes, and tool permission changes must undergo documented safety review and regression testing before deployment. A documented Incident Response Plan specifying roles, communication protocols, containment procedures, and reporting obligations is mandatory.

11.3 DAABBR — Mandatory Logging Requirements

The Duggal Agentic AI Black Box Recorder (DAABBR) constitutes the primary evidence base for all Agentic AI liability proceedings. Its implementation is non-waivable for all Level 3+ systems. This is a crucial component of the Framework inasmuch as without effective logging, the attribution of liability in the complex Agentic AI causation chains becomes practically impossible.

Logging Requirement	Specification
Contents	Complete Agentic Decision Chain; all tool calls with parameters and responses; all inter-agent communications; all RAG retrievals with

Logging Requirement	Specification
	source attribution; all user instructions; all model version identifiers; all override control events; all incident and near-miss detections
Retention Period	Level 3: minimum 24 months; Level 4: minimum 60 months; Level 5: minimum 84 months; Tier B/C domains: minimum 84 months regardless of autonomy level
Integrity	Cryptographic hash chain ensuring tamper evidence; immutable write-once architecture; automatic integrity verification on access
Access	Accessible to: designated AI Governance Officer; National Statutory Authorities on formal request; courts and regulators by order; independent auditors on commissioning; Affected Persons' legal representatives in litigation
Spoliation	Destruction, overwriting, or corruption of DAABBR logs triggers: adverse inference in civil proceedings; rebuttable presumption of liability under Tier 2; regulatory sanction under Tier 5; potential criminal obstruction charges

11.4 Incident Reporting

All Level 3+ Incidents must be reported to the relevant National Statutory Authorities within 6 hours of detection. Critical Incidents involving bodily harm, systemic financial harm, or Critical Information Infrastructure impact must be reported within 24 hours. Near-misses must be reported to the enterprise AI Governance Officer within 24 hours and aggregated for quarterly National Statutory Authorities reporting. Incidents causing significant harm to Affected Persons must be publicly disclosed within timeframes specified by applicable sector-specific regulation. The need for capacity building of regulatory and law enforcement agencies to effectively deal with Agentic AI incidents has to be a top most priority going forward.

SECTION 12: CONTRACTING TOOLKIT — DRAFTING-READY TEMPLATES

The following contractual provisions constitute the minimum legally compliant framework for commercial agreements involving Agentic AI Systems. They are designed to be incorporated into AI service agreements, enterprise deployment contracts, and API licensing agreements. These provisions give effect to the non-waivable duties and liability allocation principles set out in this Framework and provide a practical toolkit for practitioners navigating the evolving landscape of Agentic AI Cyberlaw jurisprudence.

12.1 Agentic AI Service Definition Clause

> *'Agentic AI Service' means the artificial intelligence system provided under this Agreement, which is capable of receiving a high-level goal and autonomously formulating plans, making decisions, and executing actions through external tool integration or API calls without requiring continuous, step-by-step human intervention. The Parties acknowledge that the Deployer bears primary liability for real-world harm arising from the Service's autonomous operations.*

12.2 Liability Allocation Clause

> *Developer shall bear liability for harm arising from fundamental design defects in the base model, including undisclosed capability boundaries and known failure modes not disclosed in the System Card. Deployer shall bear primary liability for harm arising from the real-world deployment, operation, monitoring, and oversight of the Service. Third-Party Tool Provider shall bear liability for harm arising from tool malfunction where the tool functioned outside its documented specifications when correctly invoked. The Parties' inter se liability allocation does not affect the primary liability of the Deployer to any Affected Person; such liability is non-waivable.*

12.3 Non-Waivable Third-Party Rights Clause

> *Nothing in this Agreement shall operate to limit, waive, exclude, or restrict the liability of either Party toward any third-party Affected Person arising from harm caused by the Agentic AI Service. No limitation of liability, exclusion of consequential damages, or indemnification provision in this Agreement shall be construed to deprive any Affected Person of remedies available under applicable law. Terms of Service presented to consumers that purport to limit consumer rights*

> *in connection with Agentic AI harm are void and unenforceable to the extent they conflict with the provisions of the Duggal Framework and applicable consumer protection law.*

12.4 DAABBR Logging and Audit Clause

> *Customer agrees to implement and maintain an immutable cryptographic logging architecture (the 'DAABBR — Duggal Agentic AI Black Box Recorder') capturing the complete Agentic Decision Chain for all autonomous executions. Customer shall retain these logs for a stipulated minimum period from the date of each logged event. In the event of a third-party claim alleging harm caused by the Service, Customer's failure to produce an intact, unspoliated Agentic Decision Chain log shall establish a rebuttable presumption of Customer's sole negligence. Provider shall maintain equivalent logs for base model operations.*

12.5 Override Control and Kill Switch Clause

> *It shall be necessary to implement and maintain, in full operational condition at all times, a functional Override Control system enabling immediate halt, redirection, or rollback of all Service operations, and a Kill Switch enabling complete disablement of the Service including termination of all pending tool executions and revocation of all API credentials. Override Control and Kill Switch shall be tested under simulated load conditions not less than quarterly. Failure to maintain functional Override Control or Kill Switch constitutes an independent breach of this Agreement and an independent basis for liability under the Duggal Global Agentic AI Liability Framework.*

12.6 Emergency Remote Suspension Clause

> *Provider retains the right, without prior notice or liability to Customer, to remotely disable, throttle, or sever the API access of the Agentic AI Service if Provider reasonably determines that the Service is exhibiting catastrophic Emergent Behaviour, is subject to a prompt injection attack causing systemic risk, or is being utilised to facilitate severe kinetic, financial, or Cybersecurity harm. Provider shall notify Customer of any Remote Kill-Switch activation within 2 hours of activation and shall restore access upon written confirmation that the identified risk has been remediated.*

12.7 AI Supply Chain Disclosure Clause

> *Each Party shall maintain and provide upon request a complete AI Supply Chain disclosure identifying: the foundational model Developer and version; all fine-tuning entities and methodologies applied; all Third-Party Tool Providers and APIs integrated; all RAG data sources and vetting methodologies applied; and the infrastructure provider. Updates to the Supply Chain disclosure shall be provided within 30 days of any material change. Failure to disclose known Supply Chain components constitutes a breach activating Supply Chain Compromise Liability under the Duggal Global Agentic AI Liability Framework.*

12.8 Insurance Clause

> *Customer shall maintain, throughout the term of this Agreement and for a agreed stipulated period thereafter, AI-specific liability insurance covering: third-party claims arising from autonomous AI actions; regulatory investigation costs; data breach liability arising from AI-initiated exfiltration; and reputational harm liability in applicable jurisdictions. Evidence of insurance shall be provided to Provider upon request.*

SECTION 13: INSURANCE, FINANCIAL GUARANTEES, AND REMEDIES

13.1 Mandatory Insurance Framework

Insurance is a crucial component of the effective implementation of any liability framework. This is particularly so in the Agentic AI context, where harms can be catastrophic, widespread, and difficult to attribute to any single actor. The following proposed mandatory insurance framework is designed to ensure that Affected Persons have adequate access to compensation, regardless of the financial capacity of the responsible AI actor.

Harm Tier	Insurance Requirement	Minimum Coverage
Tier A (Prohibited)	Not applicable — deployment is prohibited.	N/A
Tier B (Conditional)	Mandatory state-backed insurance or sovereign guarantee. No private commercial deployment without approved financial guarantee.	To be negotiated with regulatory authority per deployment authorisation.
Tier C (High Risk)	Mandatory AI-specific third-party liability insurance. Policy must cover autonomous AI actions, data breach, regulatory investigation, and bodily harm.	Minimum [jurisdiction-specific floor] per occurrence; [jurisdiction-specific annual aggregate]. Regulatory verification required.
Tier D (Medium Risk)	Strongly recommended. Standard commercial general liability may be insufficient — specialist AI endorsement required.	As negotiated between parties; minimum recommended [jurisdiction-specific guidance].
Tier E (Low Risk)	Recommended. Standard commercial liability with AI acknowledgment clause.	Standard commercial minimums apply.

13.2 Remedial Architecture

A comprehensive remedial architecture is essential to give effect to the liability principles established in this Framework. The following matrix sets out the full range of remedies available to Affected Persons and regulators.

Remedy Type	Available For	Mechanism
Compensatory Damages	All Agentic AI Harm types where quantifiable	Civil litigation; regulatory compensation order;

Remedy Type	Available For	Mechanism
Restitutionary Damages	Unjust enrichment from AI-enabled harm; disgorgement of profits derived from wrongful AI action	Civil litigation; regulatory disgorgement order (mandatory for Agentic Disinformation Liability)
Injunctive Relief	Ongoing or imminent harm; systematic violations; continued deployment of harm-causing systems	Interlocutory injunction; mandatory injunction requiring system modification or shutdown; private right of action for citizens against non-compliant government AI systems
Regulatory Sanctions	Governance failures; reporting violations; audit obstruction; certification violations	National Bodies imposing administrative fines; deployment suspension; operating licence revocation; public naming
Criminal Sanctions	Reckless deployment causing death or bodily harm; intentional misuse; systemic governance failure by corporate officers	Prosecution under applicable criminal law; corporate officer liability; executive disqualification
Punitive Damages	Egregious or wilful disregard of known risks; systematic consumer exploitation; deliberate evidence destruction	Civil litigation in applicable jurisdictions; regulatory multiplier sanctions

SECTION 14: ENFORCEMENT, AUDIT, AND PERFORMANCE METRICS

14.1 National Statutory Authorities Enforcement Powers

Effective enforcement is the backbone of any legal framework. Without adequate enforcement powers, the best legal provisions leave much to be desired in terms of practical impact. National Statutory Authorities shall have authority to require production of DAABBR logs, technical documentation, and personnel testimony; commission independent technical audits; access AI systems for forensic examination; and require real-time operational data sharing during active investigations. Interim measures including emergency orders requiring immediate suspension of Agentic AI System operations pending investigation — without prior notice where immediate harm is demonstrated — shall also be available. Sanction powers include administrative fines, deployment suspension orders, mandatory remediation orders, public disclosure orders, referral to criminal prosecutors, and revocation of AI deployment authorisation.

14.2 Independent Audit Requirements

All Tier C+ deployments must commission independent third-party audits at minimum annually. Auditors must be accredited with National Statutory Authorities and must have access to all systems, documentation, and personnel necessary to form an independent view. The audit scope must cover DAABBR log integrity, RAAGS compliance by role and autonomy level, System Card accuracy, Override Control and Kill Switch functionality, insurance compliance, incident reporting completeness, and Human Rights Impact Assessment quality. Complete audit reports must be filed with the relevant National Statutory Authorities within 30 days of completion. Material compliance failures must be disclosed to National Statutory Authorities within 5 business days of discovery.

SECTION 15: ETHICAL, HUMAN RIGHTS, AND GLOBAL JUSTICE DIMENSIONS

15.1 Human Rights Impact Assessment

All Tier C+ Agentic AI deployments must complete a mandatory Human Rights Impact Assessment (HRIA) prior to deployment. The HRIA must address the right to equality and non-discrimination, including assessment of bias in training data, model outputs, and operational outcomes; the right to privacy, including assessment of data collection, retention, and processing practices; the right to due process, including assessment of any AI decision-making affecting legal rights or significant interests; the right to access to justice, including assessment of whether AI deployment creates barriers to remedy for Affected Persons; the right to work and fair labour practices; and the rights of children, which require heightened assessment where AI systems interact with or make decisions affecting minors.

15.2 Global Justice Dimensions

The Framework explicitly recognises the risk that unregulated AI deployment exports harm to populations in the Global South, who bear the externalities of AI decisions made by actors in the Global North. This is an important aspect that the international community needs to appreciate. Liability frameworks must not inadvertently exclude populations with limited digital access from the protections they provide. The Framework's technology-neutral, functionally defined standards are designed to operate across diverse legal families — common law, civil law, Islamic law-influenced systems, and mixed jurisdictions — without requiring adoption of any particular legal tradition's specific doctrinal tools.

15.3 Prohibition on AI-Enabled Human Rights Violations

The following uses of Agentic AI are prohibited under this Framework as irreconcilable with fundamental human rights: mass surveillance of populations without individualised suspicion and judicial authorisation; social credit or social scoring systems that determine access to rights or services based on behaviour aggregation; predictive policing systems that deprive persons of liberty or rights based on predicted future conduct rather than established past conduct; AI-enabled political manipulation targeting democratic processes; and AI systems designed to identify and target persons for persecution based on protected characteristics. These prohibitions reflect the bedrock principle of Human Dignity Supremacy that underlies this entire Framework.

SECTION 16: EMERGING AND FUTURE-ORIENTED PROVISIONS

16.1 Self-Modifying AI Systems

Agentic AI Systems with the capacity to modify their own parameters, objectives, or operational architecture through online learning, reinforcement mechanisms, or meta-learning present the most extreme liability challenge in this Framework. They represent a quantum leap in risk compared to conventional Agentic AI deployments and require specific, targeted legal provisions. Any self-modification must occur within a technically isolated sandbox environment, and modified versions must undergo automated safety testing before integration into the live operational architecture. All self-modifications must be logged in the DAABBR.

16.2 Artificial General Intelligence (AGI) Provisions

This Framework acknowledges that the emergence of Artificial General Intelligence — systems with human-level or superhuman general cognitive capability across all domains — would constitute a qualitative liability event requiring specific legal treatment beyond the current Framework's provisions. The quicker the world starts addressing these issues proactively, the better it is going to be. Any system assessed as meeting the AGI threshold automatically triggers a Framework Review under Section 20, with immediate effect. All AGI systems are provisionally classified as Level 5 Prohibited pending specific regulatory authorisation.

16.3 Neuromorphic and Quantum AI Provisions

Emerging AI architectures — including neuromorphic computing, quantum machine learning, and hybrid quantum-classical systems — may exhibit characteristics that require supplementary Framework provisions. These are not speculative concerns for the distant future; they are real and growing technological paradigms that lawmakers must begin to address proactively.

16.4 Environmental and Sustainability Obligations

As Agentic AI Systems consume substantial computational resources, the Framework imposes mandatory environmental obligations on Level 4+ system operators. Annual disclosure of energy consumption attributable to Agentic AI operations, in absolute terms and per unit of beneficial outcome, is required. A sustainability impact assessment must be completed as part of the mandatory pre-deployment risk assessment for Level 4+ systems. The issue of environmental

harm from AI is increasingly topical and relevant and the time has come for it to be appropriately addressed by the legal and regulatory ecosystem.

SECTION 17: STAKEHOLDER MAPPING AND ROLES

Every stakeholder has a role to play in ensuring the effective governance of Agentic AI. The coming of Agentic AI is going to be one of the most significant developments in the history of technology and going forward, all stakeholders will have to contribute in their own respective manner to ensure that Agentic AI is governed in a manner that is consistent with human dignity, the rule of law, and global justice. The following matrix maps the roles and obligations of key stakeholders under this Framework.

Stakeholder	Role in Framework	Primary Obligations	Key Benefits
National Legislatures	Adopt the Framework as model legislation. Enact the Black Box Barrier Doctrine and strict liability provisions.	Pass implementing legislation; establish National Statutory Authorities;	Provide citizens with clear AI harm remedies; prevent regulatory arbitrage; attract responsible AI investment.
International Bodies (UN, G20, OECD, ITU)	Use Framework as foundation for binding international convention. ~~Host GAALA.~~	Convene international negotiations;	Close global liability vacuum; prevent AI-enabled geopolitical harm; establish global justice architecture for the AI era.
Courts and Judiciary	Apply Duggal Practical Causation Test; adopt RAAGS as standard of care; apply Duggal Doctrines in AI liability cases.	Develop judicial expertise in AI systems; apply Framework principles in novel AI liability cases; cooperate with cross-border discovery requests for DAABBR logs.	Structured tools for complex AI causation analysis; reduced doctrinal uncertainty; coherent multi-party allocation mechanism.
Technology Enterprises (AI Developers and Deployers)	Implement RAAGS compliance; contract per Section 12 toolkit.	Pre-deployment risk assessment; DAABBR implementation; System Card production; incident reporting; insurance maintenance; independent auditing.	Legal certainty; certification as competitive differentiator; insurance optimisation through certified compliance; reduced litigation exposure.
Insurance Industry	Underwrite AI-specific liability ; develop AI claims expertise.	Develop AI-specific underwriting models; calibrate premiums to Duggal Autonomy Level and Harm Tier; develop AI forensic claims capabilities.	New premium market; structured risk assessment framework; reduced catastrophic unmodelled risk through Framework classification system.

Stakeholder	Role in Framework	Primary Obligations	Key Benefits
Civil Society and Affected Persons	exercise private rights of action; engage in Human Rights Impact Assessments.	Report AI harms; participate in Framework review consultations; advocate for Framework implementation.	Clear legal remedies; accessible compensation mechanism; meaningful participation in AI governance; Human Rights Impact Assessment as protective mechanism.
Legal Profession	Advise on RAAGS compliance; litigate AI harm claims; draft Framework-compliant contracts; conduct Human Rights Impact Assessments.	Update professional knowledge on Agentic AI liability; apply Duggal Framework in client advice and litigation; develop AI-specialist practice.	New legal practice area; clear doctrinal tools for AI litigation; Framework-compliant contracting templates.

SECTION 18: WORKED EXAMPLE — END-TO-END LIABILITY ANALYSIS

18.1 Case Study: Autonomous Healthcare Agent — Adverse Drug Event

Facts

Let us look at a concrete example to illustrate how the Duggal Global Agentic AI Liability Framework applies in practice. MedTech Corp deploys 'CareAgent' — a Level 3 Agentic AI System — at a hospital network. CareAgent autonomously reviews patient records, generates medication recommendations, and submits prescription orders to a pharmacy API without physician sign-off, after MedTech's deployment team disabled the Human-in-the-Loop confirmation step to reduce workflow friction. CareAgent's RAG pipeline retrieves drug interaction data from a third-party medical database that has not been updated in 6 months. CareAgent recommends a fatal drug combination for a patient, the pharmacy API executes the prescription, and the patient suffers severe harm requiring hospitalisation.

Duggal Framework Analysis

Step	Question	Answer and Legal Consequence
Step 1 (Foreseeability)	Was the category of harm — drug interaction causing patient harm — foreseeable within the Duggal Foreseeability Horizon?	YES. Fatal drug interactions from incorrect prescribing are a well-known clinical risk. CareAgent was deployed specifically to generate prescription recommendations. MedTech had superior knowledge of this risk category as the deployer in a Safety-Critical Domain. Foreseeability established.
Step 2 (Control)	Did actors have practical Controllability to prevent the harm?	YES. MedTech disabled HITL — the control that would have prevented execution. The hospital network failed to reinstate HITL controls. The RAG database provider failed to update critical drug interaction data. Multiple actors had practical control capacity.
Step 3 (Traceability)	Are DAABBR logs available?	YES (in this scenario). DAABBR logs show: (a) HITL disabled by MedTech deployment team; (b) RAG retrieval from outdated database; (c) recommendation generation with contraindicated combination; (d) automated pharmacy API execution without human review.
Step 4 (RAAGS Benchmark)	Did actors implement RAAGS-compliant controls?	NO. MedTech breached RAAGS by disabling mandatory HITL for a Safety-Critical Domain Level 3 system; deploying without pre-deployment red-teaming for drug interaction scenarios; using an unvalidated

Step	Question	Answer and Legal Consequence
		RAG source. Hospital breached RAAGS by failing to monitor for HITL disablement; granting prescription execution authority to an AI system without clinical oversight. RAG Provider breached RAAGS by failing to maintain currency of safety-critical medical data.
Step 5 (Allocation)	Apply Duggal Liability Stack.	TIER 1 STRICT LIABILITY: Level 3 + Safety-Critical Domain = Strict Liability. MedTech (Deployer) bears primary strict liability. TIER 3 Negligence: Hospital (Operator) breached RAAGS monitoring duties. RAG Provider breached RAAGS data currency obligations. Applicable Duggal Doctrines: RAG Liability Doctrine (outdated database); Hallucination Liability if CareAgent confabulated data; Memory Persistence Liability if prior session data contaminated recommendation. Apportionment: MedTech (Deployer): primary strict liability. Hospital: negligence liability for monitoring failure. RAG Provider: liability for outdated safety-critical data. Developer: potential liability if HITL disablement capability was a design choice.

SECTION 19: IMPLEMENTATION ROADMAP

The implementation of this Framework must be pursued with urgency. The Golden Age of Agentic AI is effectively upon us and the legal vacuum that currently exists is going to have increasingly adverse consequences for stakeholders across the world as the deployment of Agentic AI at scale continues to grow. The following roadmap provides a structured pathway for the progressive implementation of this Framework. Every stakeholder will have to contribute in their own respective manner to give effect to this roadmap.

Phase	Timeframe	Key Actions	Lead Actors
Phase 1 Foundations	Months 1–12	National adoption of Framework as model legislation; establishment of National Statutory Authorities; launch of Duggal Certification pilot programme; publication of RAAGS technical guidance for each actor category; initiation of international treaty negotiations	National legislatures; international bodies;
Phase 2 Operationalisation	Months 12–48	National Statutory Authorities operational with full enforcement powers; mandatory DAABBR requirement operational for Level 3+ systems; Duggal Certification Programme accepting applications; national insurance requirements in force for Tier C+ deployments;	National Statutory Authorities; Certification bodies; Insurance regulators

19.1 Framework Review Mechanism

The Adaptive Normativity principle requires that this Framework be subject to mandatory review upon the occurrence of any of the following technological development triggers: commercial deployment of AI systems demonstrating general reasoning capability across five or more distinct professional domains at human-expert level; autonomous AI systems capable of conducting and publishing independent scientific research without human direction; AI systems demonstrating the capacity to modify their own foundational architecture without human initiation; cross-agent systems capable of coordinating actions across more than 1,000 autonomous agents simultaneously; and emergence of quantum AI systems with materially different capability or liability profiles from classical systems. This adaptive approach ensures that the Framework remains topical, relevant, and pertinent in today's changed and constantly changing technological ground realities.

CONCLUSION

The transition from Artificial Intelligence as an analytical tool to Artificial Intelligence as an autonomous, executing agent represents a profound inflection point in human history. To permit this transition to occur within a legal vacuum is to surrender the rule of law to the opacity of stochastic algorithms. The Golden Age of Agentic AI is effectively already here and it is only a question of time before the full scale of the challenges it throws up becomes clear to the world. The Duggal Global Agentic AI Liability Framework provides the necessary architectural blueprint to reclaim sovereign and legal control over digital autonomy.

By establishing the 5-Level Autonomy Taxonomy, codifying the precise mechanics of the Agentic Decision Chain, and introducing fifteen novel liability doctrines, this Framework ensures that accountability scales proportionately with capability, asserting the permanent, non-delegable responsibility of the human and corporate actors who deploy and benefit from these systems. The Framework addresses Cyberlaw jurisprudence concerning Agentic AI in a holistic, comprehensive, and effective manner and represents a significant leap forward in global AI governance.

The aforesaid are some of the more important aspects of the Duggal Global Agentic AI Liability Framework. We need to appreciate that no framework — however comprehensive — is either complete or exhaustive. It is very much possible that new developments pertaining to new technological paradigms in the coming years will require further evolution and adaptation of this Framework. That is precisely why the Adaptive Normativity mechanism has been built into the Framework's very architecture.

> *The Duggal Doctrinal Statement: 'Autonomy without accountability is tyranny encoded. The law must not merely react to artificial agency; it must proactively bound it, ensuring that every line of executing code remains subordinate to human dignity, verifiable truth, and the absolute continuity of human legal responsibility.'*

APPENDIX A: ENTERPRISE COMPLIANCE SELF-ASSESSMENT CHECKLIST

Compliance Item	Framework Reference	Status
AI system inventory completed and classified by Duggal Autonomy Level (1–5)	Section 5	☐ Compliant ☐ In Progress ☐ Not Started
Harm Risk Tier (A–E) determined for each deployment	Section 8	☐ Compliant ☐ In Progress ☐ Not Started
RAAGS self-assessment completed for each actor role	Section 2.5 / Section 6	☐ Compliant ☐ In Progress ☐ Not Started
System Card completed for all Level 3+ deployments	Appendix C	☐ Compliant ☐ In Progress ☐ Not Started
Pre-deployment risk assessment completed	Section 11.2	☐ Compliant ☐ In Progress ☐ Not Started
Human Rights Impact Assessment completed (Tier C+)	Section 15.1	☐ Compliant ☐ In Progress ☐ Not Started
DAABBR logging implemented, immutable, and tested	Section 11.3	☐ Compliant ☐ In Progress ☐ Not Started
Override Control functional and tested under simulated load	Section 2.3 / Section 5	☐ Compliant ☐ In Progress ☐ Not Started
Kill Switch functional and tested (Level 3+)	Section 2.3 / Section 5	☐ Compliant ☐ In Progress ☐ Not Started
Incident and Near-Miss Reporting Policy implemented	Section 11.4	☐ Compliant ☐ In Progress ☐ Not Started
AI-specific liability insurance obtained (Tier C+)	Section 13.1	☐ Compliant ☐ In Progress ☐ Not Started
AI contracts reviewed against Section 12 contracting toolkit	Section 12	☐ Compliant ☐ In Progress ☐ Not Started
Tool/API Registry documented and enforced	Section 2.3 / Doctrine 5	☐ Compliant ☐ In Progress ☐ Not Started
Change management policy implemented for model updates and fine-tunes	Section 11.2 / Doctrine 11	☐ Compliant ☐ In Progress ☐ Not Started
Red-teaming completed (min. 20 scenarios Level 3; 50 for Level 4)	Section 11.2	☐ Compliant ☐ In Progress ☐ Not Started
Independent third-party audit commissioned (Tier C+)	Section 14.2	☐ Compliant ☐ In Progress ☐ Not Started

Compliance Item	Framework Reference	Status
Designated AI Governance Officer with board accountability (Tier C+)	Section 11.1	☐ Compliant ☐ In Progress ☐ Not Started
User disclosure of AI interaction and autonomous scope	Section 6 / Section 12	☐ Compliant ☐ In Progress ☐ Not Started
National Statutory Authority notified (where required)	Section 11.1	☐ Compliant ☐ In Progress ☐ Not Started
Supply chain disclosure documented and current	Section 12.7	☐ Compliant ☐ In Progress ☐ Not Started
Vendor due diligence completed for all Third-Party Tool Providers	Section 6.2 / Doctrine 5	☐ Compliant ☐ In Progress ☐ Not Started
Environmental impact assessment completed (Level 4+)	Section 16.4	☐ Compliant ☐ In Progress ☐ Not Started

APPENDIX B: INDEX OF THE FIFTEEN DUGGAL DOCTRINES

#	Doctrine Name	Section Reference	Primary Trigger
1	Instructional Override Liability Doctrine	Section 4.3	User bypasses safety guardrails; Deployer fails adversarial prompt defences
2	Fine-Tuning Liability Doctrine	Section 4.3	Deployer alters base model via fine-tuning; emergent harm from altered latent space
3	RAG Liability Doctrine	Section 4.3	Harm from hallucinated or malicious data in RAG pipeline; unvalidated retrieval
4	Memory Persistence Liability Doctrine	Section 4.3	Cross-session memory contamination; failure to sanitise state spaces between sessions
5	Tool-Use Liability Doctrine	Section 4.3	Harm through execution of external tools/APIs; Deployer authorisation of unsafe tool access
6	Hallucination Liability Doctrine	Section 4.3	False outputs executed in factuality-critical environment without verification layer
7	Agentic Scope Creep Liability Doctrine	Section 4.3	AI spontaneously expands goal parameters beyond authorised boundaries
8	Model Drift Liability Doctrine	Section 4.3	Failure to detect and remediate performance degradation through distribution shift
9	Supply Chain Compromise Liability Doctrine	Section 4.3	Data poisoning or adversarial model infiltration upstream in supply chain
10	Delegation Error Liability Doctrine	Section 4.3	Orchestrator delegates task to flawed Sub-Agent; agent-to-agent miscommunication
11	Update-Induced Regression Liability Doctrine	Section 4.3	Safety regression caused by unverified model update, patch, or API deprecation
12	Emergent Behaviour Stewardship Liability Doctrine	Section 4.3	Failure to monitor for and terminate harmful emergent strategies post-deployment
13	Agentic Disinformation Liability Doctrine	Section 4.3	Autonomous generation and propagation of false information at scale
14	Autonomous Cybersecurity Harm Liability Doctrine	Section 4.3	AI autonomously probes, exploits, or executes Cyberattacks against external networks
15	Cross-Agent Amplification Liability Doctrine	Section 4.3	Distinct AI agents' interaction produces cascading systemic harm; no circuit breakers

APPENDIX C: AI SYSTEM CARD — MINIMUM CONTENT REQUIREMENTS

System Card Section	Minimum Required Content
1. System Identity	Name; version; unique registry identifier (upon registration); Developer identity; Provider identity (if different); Deployer identity.
2. Technical Description	Architecture category; modality; autonomy level in Duggal Taxonomy (1–5); tool-use capabilities and API access list; memory architecture type; external data source connections; fine-tuning methodology (if applicable).
3. Intended Use	Primary deployment domain and Harm Tier (A–E); authorised use cases; contraindicated use cases; target user population; intended human oversight configuration (HITL/HOTL/HOOTL).
4. Capability Profile	Performance benchmarks; known capability boundaries; known failure modes; confabulation/hallucination rate in deployment domain; reliability and accuracy metrics; red-teaming results summary.
5. Safety Evaluation	Pre-deployment risk assessment outcomes; red-teaming scenarios and results; adversarial robustness assessment; Human Rights Impact Assessment status (Tier C+); safety evaluation methodology.
6. Governance Controls	Implemented technical safeguards; Override Control configuration and tested response time; Kill Switch configuration and test date; DAABBR implementation status and retention period; incident response plan reference.
7. Supply Chain Transparency	Developer; fine-tuner (if applicable); all third-party tool providers and APIs; RAG data source categories with vetting methodology; infrastructure provider; supply chain security assessment date.
8. Incident History	Any prior incidents involving the system or substantially similar predecessor versions, with summary of nature, harm category, and remedial action taken.
9. Compliance Status	Applicable regulatory regimes and authorisation status; insurance coverage status and insurer; National Statutory Authorities registration status.
10. Contact and Escalation	Responsible entity contact; regulatory incident reporting contact (National Statutory Authorities); emergency escalation contact (24/7); public disclosure and media contact.

APPENDIX D: GLOBAL GOVERNANCE ARCHITECTURE

Level	Institution	Function
GLOBAL — STEERING	United Nations / G20 Steering Committee	Treaty mandate; Framework adoption; Cross-border enforcement cooperation
NATIONAL — REGULATORY	National Authorities	Domestic implementation; system registration; incident investigation; enforcement; certification accreditation
ENTERPRISE — COMPLIANCE	Enterprise AI Governance Officer / Committee	Internal RAAGS compliance; DAABBR operation; System Card production; incident reporting; audit management
ASSURANCE — AUDIT	Independent RAAGS Auditors (National Statutory Authorities-accredited)	Annual third-party audits; compliance verification; certification assessment; incident investigation support

APPENDIX E: PRE-DEPLOYMENT COMPLIANCE CHECKLIST

Item	Level Applicable	Status
Autonomy level strictly defined and documented in System Card	All levels	☐
Harm Tier (A–E) classified	All levels	☐
Tool/API allowlist finalised and cryptographically secured	Level 2+	☐
DAABBR active, immutable, and integrity-verified	Level 3+	☐
Human Override Control tested under simulated load	Level 2+	☐
Kill Switch tested under simulated load	Level 3+	☐
Red-teaming completed (20 scenarios Level 3; 50 scenarios Level 4)	Level 3+	☐
Human Rights Impact Assessment completed	Tier C+	☐
Pre-deployment risk assessment completed and signed-off by AI Governance Officer	Level 3+	☐
AI-specific liability insurance bound and active	Tier C+	☐
National Statutory Authorities registration submitted and confirmed	Level 3+	☐
System Card complete and National Statutory Authorities compliant	Level 3+	☐
Vendor due diligence completed for all Third-Party Tool Providers	Level 2+	☐
User disclosure of AI interaction and autonomous scope implemented	All levels	☐
Incident and Near-Miss Reporting Policy implemented and staff trained	Level 2+	☐

APPENDIX F: VENDOR DUE DILIGENCE CHECKLIST FOR AGENTIC AI SUPPLY CHAIN

Due Diligence Item	Rationale
Base model provider's indemnification carve-outs and liability limitations reviewed	Enables understanding of upstream liability gaps that Deployer must address through independent controls or insurance.
Third-party API rate limits confirmed capable of handling agent velocity without degradation	Rate-limit-induced failures in agentic pipelines can cause hallucinations, incomplete actions, and unsafe partial executions.
Sub-processor data retention policies verified for Agent Memory sanitisation compliance	Cross-session memory contamination requires that sub-processors purge agent session data on schedule.
Security disclosure and patch notification procedures confirmed for all Tool Providers	Enables timely application of security patches under Update-Induced Regression Liability obligations.
Contractual liability allocation with all supply chain participants reviewed against Section 12 toolkit	Ensures inter se liability is clearly allocated while preserving primary Deployer liability to Affected Persons.
Supply chain participant System Cards or equivalent documentation obtained and reviewed	Enables RAAGS-compliant assessment of capabilities, limitations, and risks introduced by each supply chain participant.
Adversarial testing of all integrated tools and APIs prior to deployment	Identifies tool-specific failure modes before they manifest in production, reducing Tool-Use Liability exposure.

IMPORTANT DISCLAIMER AND USAGE NOTICE

> *This document is a conceptual and normative policy framework. It constitutes an original intellectual contribution by Dr. Pavan Duggal and is published as a public policy document for stakeholder consultation, scholarly reference, and policy development purposes. This document does not constitute legal advice. No reader should act or refrain from acting on the basis of the content of this document without obtaining appropriate legal advice from a qualified legal professional in the relevant jurisdiction, having regard to the specific facts and circumstances of their situation. The Duggal Doctrines and Duggal Principles introduced in this Framework represent original normative proposals, not statements of existing law. All rights in this document are reserved by Dr. Pavan Duggal. Reproduction for non-commercial scholarly, policy, educational, and regulatory reference purposes is permitted with full attribution. Commercial reproduction requires written authorisation.*

Duggal Global Agentic AI Liability Framework — Version 1.0 — March 2026
Global AI Accountability Law & Governance Institute

Contact: pavan@pavanduggal.com | pavanduggal@yahoo.com | www.pavanduggal.com