INDIAN GOVERNANCE AND POLICY PROJECT

AI Privilege: Emerging Questions of Legal Protection

AI PRIVILEGE: EMERGING QUESTIONS OF LEGAL PROTECTION

October 2025

Published by
Indian Governance and Policy Project (IGAP)

Authors
Soumya AK and Ananya Agrawal

Designer
Manoj Murali

About IGAP

The Indian Governance and Policy Project (IGAP) is a policy, business advisory, and research studio working at the intersection of governance, technology, markets, and national development.

Grounded in a clear understanding of how state capacity, market forces, and emerging technologies shape India's strategic trajectory, IGAP addresses key questions that define the country's future — from the governance of AI and digital infrastructure to financial innovation, sustainability, and national security.

Bringing together lawyers, policy thinkers, and strategists with deep business and geopolitical insight, IGAP delivers solutions that balance India's developmental and security priorities with its democratic values and constitutional principles.

Table of Contents

01	The Catalyst	01
02	What Is AI Privilege?	01
03	Understanding Traditional Privilege	01
04	Elements-Based Analysis: Could AI Privilege Meet Traditional Standards?	03
05	The Case For AI Privilege	05
06	Doctor-Patient Confidentiality	06
07	The Case Against AI Privilege	07
08	Way Forward	09

The Catalyst

In June 2025, The New York Times obtained a court order compelling OpenAI to preserve all user chat logs, including temporary and deleted conversations, as part of ongoing copyright litigation. OpenAI CEO Sam Altman publicly called this "an inappropriate request that sets a bad precedent,"[1] arguing that the same level of protection afforded to conversations with human professionals should extend to AI interactions.

While millions of people worldwide are turning to AI chatbots for some of their most sensitive questions and intimate concerns, these conversations enjoy none of the legal protections afforded to communications with lawyers, doctors, or psycho-therapists.

Are Your E-Signed Contracts Privileged?

The Indian market for contract-drafting and e-signature platforms is fundamentally grounded in confidentiality, data protection, and statutory e-signature validity under the Information Technology Act, 2000. Popular platforms such as Zoho Sign, eMudhra, Spotdraft, DocuSign, Adobe Sign ensure secure and compliant execution of agreements through encryption, audit trails, but they do not extend the legal protections of privilege, which arise solely from the advocate–client fiduciary relationship.

What Is AI Privilege?

While there is no universally agreed definition, *"AI privilege"* refers to a proposed legal framework that would protect communications between users and artificial intelligence systems from compelled disclosure in legal proceedings.[2] Similar to how attorney-client privilege shields confidential legal consultations, AI privilege would create a zone of confidentiality around human-AI interactions.

Currently, this privilege does not exist in any jurisdiction.

Understanding Traditional Privilege

To evaluate whether AI privilege makes sense, we must first understand what makes a communication "privileged" under existing law.

Attorney-Client Privilege

The doctrine of privileged communication traces its origins to English common law, where courts recognised that compelling a legal adviser to disclose confidential communications with a client would undermine the administration of justice. Clients, fearing exposure, might otherwise withhold information essential for effective representation. By the 17th and 18th centuries, English courts had affirmed that the privilege belongs to the client, not the lawyer, and cannot be waived without the client's consent. This classical formulation of privilege was subsequently adopted across common law jurisdictions such as the United States, Canada, Australia, and India.

Common law jurisdictions recognise it as a substantive right, while civil law systems typically treat it as a procedural safeguard. Its underlying rationale is to foster open and honest communication between clients and their legal advisers. Consequently, it is excluded from evidence to uphold the integrity of legal advice and the justice process.

What AI Privilege Is NOT

Data Privacy

Confidentiality Agreements

Trade Secret Protection

Essential Elements:

1. A communication
2. Between privileged persons (attorney and client)
3. Made in confidence
4. For the purpose of seeking, obtaining, or providing legal assistance

Legal Foundation in India: Sections 132-134 of the Bharatiya Sakshya Adhiniyam, 2023 (BSA) establish the common law principles governing professional communications between attorneys and clients. In addition, Rule 7 of the BCI's Rules on an Advocate's Duty Towards the Client, reinforces this framework by mandating that an advocate must maintain the confidentiality of all communications with the client.

Elements-Based Analysis: Could AI Privilege Meet Traditional Standards?

Element 1:
A Communication

Traditional Standard: An exchange of information between parties.

AI Context: ⭕ Users communicate with AI systems through text, voice, or other inputs.

> ? *Is AI a "party" capable of receiving communications? What if the communication is stored, analyzed, or used to train other AI?*

Element 2:
Between Privileged Persons

Traditional Standard: A professional relationship with defined duties and qualifications.

AI Context: ❌ No fiduciary duty exists as the relationship is commercial and AI is not a licensed professional. Further, AI cannot make independent decisions about waiving privilege. Multiple parties (AI company, user, service providers) undermine confidentiality

> ? *Can a non-human, non-professional entity be a "privileged person"?*

Work Product Doctrine

Under U.S. law, the work-product doctrine protects any documents and tangible materials prepared in anticipation of litigation.[4] Attorney mental impressions, conclusions, and legal theories (opinion work product) are nearly always protected, while factual materials like witness notes or chronologies (fact work product) are discoverable only upon substantial need and undue hardship.

In India, there is no statutory doctrine, but courts[5] have recognised similar protection.

? Should inputs to AI tools be treated as extensions of legal work products, or as third-party disclosures?

? Should courts treat materials generated or refined through AI as privileged work product or discoverable evidence?

Element 3:
Made in Confidence

Traditional Standard: Communication intended to remain private, with reasonable precautions taken.

AI Context: 👱 Users may intend confidentiality and Terms of service often have clauses on privacy [3]

❌ But AI companies retain access to conversations and Data is often used for training or improvement. Additionally, multiple parties have technical access and employees (human oversight) may review conversations.

? *Can there be confidence when a corporation retains indefinite access to communications for its own purposes?*

Element 4:
For the Purpose of Seeking/Obtaining Professional Assistance

Traditional Standard: The communication must relate to securing professional services.

AI Context: Users do seek advice, information, or assistance.

❌ But the purpose is not always clear, Users often don't seek "professional" help, just convenient answers. AI provides information, not professional services. The purpose may be recreational, commercial, or unclear

? *Can AI offer professional assistance (ties in with Element 1)*

The Case For AI Privilege

Argument 1

User Trends: Evolving Patterns of Trust and Reliance

Millions of individuals now turn to artificial intelligence systems for sensitive conversations — mental health counseling, preliminary medical assessments, legal information gathering, financial planning, and intimate personal guidance. This behavioural shift reflects not merely convenience but often necessity, as users seek alternatives to traditional professional services that may be unavailable or resource intensive.

Why this matters

Users may have reasonable expectations that sensitive conversations are private. AI providers often market their services as confidential spaces. The effectiveness of AI assistance may depend on user candor.

Argument 2

Encourages Disclosure

Traditional professional privileges exist to encourage disclosure. Attorney-client privilege promotes effective legal representation by ensuring clients can speak candidly. Doctor-patient confidentiality facilitates accurate diagnosis by enabling patients to disclose symptoms they might otherwise conceal.

Why this matters

AI legal guidance tools could improve access to justice if individuals are not deterred from seeking information due to concerns about future legal consequences. Public health systems could benefit from candid disclosure to medical AI diagnostic tools, enabling early intervention. The absence of privilege protection may create a chilling effect that undermines these potential societal benefits.

Doctor-Patient Confidentiality

Unlike attorney–client privilege, doctor–patient confidentiality in India is primarily ethical and professional and arises under multiple frameworks that collectively impose a duty of confidentiality. However, this duty is not absolute. Disclosure is permitted in limited circumstances, including when required by law or court order, or where necessary to address risks to life, health, public safety.[6]

01 A confidential communication

02 Between patient and physician

03 Made in the course of medical treatment

Argument 3

Equality and Access

Traditional professional privileges disproportionately protect individuals with sufficient economic resources to retain attorneys and physicians. Artificial intelligence has demonstrated potential to dramatically reduce barriers to accessing professional guidance, yet denying privilege protection to AI communications risks creating a two-tiered confidentiality regime stratified by economic class.

Why this matters

As artificial intelligence tools increasingly serve as the primary mechanism through which economically disadvantaged populations access legal, medical, and mental health guidance, the absence of privilege protection effectively denies confidentiality protections to those already most vulnerable. Expanding privilege doctrine to AI communications could democratize confidentiality protections traditionally available only to the affluent.

Argument 4

Technological Evolution Requires Legal Adaptation

Legal privilege doctrine has historically evolved in response to technological change. When attorney-client privilege was initially recognized, modern communication technologies (telephones, email, videoconferencing) did not exist. Courts extended privilege protection to these new modalities.

Why this matters

Failing to consider whether AI communications warrant similar protection risks allowing privilege doctrine to ossify, protecting only historical professional relationships while leaving modern equivalents unprotected, thereby undermining the very policies that justify privilege recognition.

The Case Against AI Privilege

Argument 1

The Absence of Professional Judgment and Ethical Duties

Lawyers and doctors are bound by professional ethics, licensing requirements, and personal liability. They exercise independent judgment and can be disciplined or sued for misconduct. These professional obligations are not merely incidental features but are central to why privilege protection is granted to certain communications.

Why this matters

AI has no professional license that can be revoked, cannot be held personally liable for bad advice, and is not bound by codes of professional conduct. AI cannot exercise independent judgment or recognize when it should refuse to help. AI providers' interests may conflict with users' interests. Privilege protects professional relationships precisely because professionals have enforceable duties to their clients and patients. Without ethical constraints, privilege could be abused. AI cannot claim privilege on behalf of users or help them assert it. The relationship is fundamentally commercial, not professional.

Argument 2

Technical and Practical Problems

AI systems are owned by private companies, trained on public data, and often involve multiple parties in the processing chain.

Why this matters

AI conversations often involve the user, the AI company, cloud service providers, and potential third-party processors. Privilege traditionally requires confidentiality limited to the privileged parties. AI companies use conversations to improve their models. This ongoing access to content is incompatible with traditional privilege concepts. Users cannot ensure conversations remain private. The AI company controls data retention, employee access, and response to legal process. It is unclear as to what counts as an AI privileged communication.

Argument 3

Interference with Legitimate Investigations

Privilege claims can obstruct criminal investigations, civil litigation, and regulatory oversight. Creating a new category of privilege protection has significant consequences for law enforcement and the justice system.

Why this matters

Unlike lawyers and doctors, AI cannot judge whether disclosure is legally required. The breadth of AI use means vast amounts of potentially relevant evidence could be privileged. In civil litigation, parties could hide damaging evidence by routing it through AI systems. Discovery is already difficult; AI privilege would complicate it further. Unlike professional privileges with defined scope, AI privilege could encompass almost any topic. Courts and investigators would face new barriers to obtaining evidence necessary for justice.

Argument 4

Lack of Social Necessity

Given the balancing of interests, arguing it on grounds of access and social necessity is not a compelling reason. Society recognizes professional privileges because the alternative is intolerable, we need people to get legal representation and medical care.

Why this matters

People used AI before any privilege discussion, suggesting privilege isn't needed to encourage use. AI is a commercial product, not a professional necessity. Unlike lawyers and doctors, AI alternatives are easily available. Creating privilege to protect a commercial product sets a troubling precedent. This also raises boundary problems: What about conversations with customer service chatbots? Do navigation apps get privilege for location data? Could any digital service claim privilege? Without clear answers, AI privilege could open the door to unlimited expansion of privilege claims across the technology sector.

Way Forward

What qualifies as "AI" for privilege purposes?

- Only conversational AI like ChatGPT?
- Specialized AI (medical diagnosis, legal research)?
- All AI systems?

Who owns the privilege?

- The user (like attorney-client privilege)?
- The AI company (like corporate privilege)?
- Both jointly?
- Can one waive it without the other's consent?

When does a privileged AI conversation begin and end?

- From the first query?
- Only conversations explicitly marked as private?
- Does it include all interactions with an AI service?
- What about conversations that include non-privileged content?

What is the nature of the relationship between humans and AI?

Is it:

- A tool relationship (like using a calculator or search engine)?
- A service relationship (like consulting a professional)?
- Something entirely new that doesn't fit existing categories?

The emerging debate on AI privilege raises unresolved conceptual and legal challenges — whether AI functions as a tool, a service, or something in between will shape how confidentiality and privilege are understood in this context. If AI is merely a tool, privilege may not arise at all. But if it increasingly performs advisory or interpretive functions, akin to professional services, the argument for extending privilege becomes more compelling yet also more complex.

As AI becomes more capable, more trusted, and more integrated into sensitive aspects of life, the question of privilege becomes not just a legal technicality but a fundamental issue about privacy, power, and the future of human-AI relationships.

Endnotes

1. https://x.com/sama/status/1930785054005076100;
In July 2025, OpenAI CEO Sam Altman appeared on podcaster Theo Von's show and warned that: "There is not currently a legal privilege that protects sensitive personal data someone shares with ChatGPT if a subpoena compels OpenAI to provide that information."

2. https://itlawco.com/ai-privilege-protecting-user-interactions-with-generative-ai/

3. See Clauses 4, 5, and 7 of OpenAI's Privacy Policy and its Security and Privacy Statement. Google's Gemini policies specify that data used to improve AI models is anonymised and detached from user identity before any human review, and that the company implements industry-standard security measures, such as encryption and access controls, to safeguard user data against unauthorised access or disclosure (Gemini API Terms; Google Privacy Policy).

4. Hickman v. Taylor, 329 U.S. 495 (1947)

5. The Bombay High Court, in Larsen & Toubro Ltd v Prime Displays, held that documents and communications prepared in anticipation of litigation are protected under Sections 126 and 129 of the Evidence Act.

6. Professional conduct rules, healthcare establishment regulations, and sectoral laws such as the Mental Healthcare Act, 2017 reinforce this obligation while allowing exceptions in defined cases. Indian Medical Council (Professional Conduct, Etiquette and Ethics) Regulations, 2002 (Reg. 7.14): bars physicians from disclosing patient secrets, with exceptions for court orders, risk to others, or notifiable diseases; Clinical Establishments (Registration and Regulation) Act, 2010: Requires every registered clinical establishment to maintain confidentiality of patient records and produce them only to the patient or an authorised authority; Mental Health Act 2017 (Section 23): Requires all health professionals providing care or treatment to a person with mental illness to keep all such information confidential which has been obtained during care or treatment with a few exceptions including harm or violence risks, court orders, public safety and security.

For more details visit: www.igap.in

relations@igap.in | igap.in