

SLTA | SWISS
LEGALTECH
ASSOCIATION

in collaboration with

P&TS
INTELLECTUAL PROPERTY

LAW & TECH INSIGHTS MAGAZINE

2026-11

SLTA TECH & LAW INSIGHTS

GENERATIVE AI AND INFORMATION SECURITY

SPECIFIC RISKS FOR SWISS LAW FIRMS AND PATENT ATTORNEYS

By Christophe Saam, CEO at P&TS SA

Editorial Team: Rucsandra Bratu

INTRODUCTION

Generative AI is already present in most law firms, whether management is aware of it or not. It brings genuine efficiency gains, but introduces risks that are qualitatively different from traditional cybersecurity threats. This chapter complements the SAV/FSA Guidelines on the Use of Artificial Intelligence of June 2024 (consolidated February 2025) from an information security perspective, and is also relevant to patent attorneys, who are subject to equivalent professional secrecy obligations. It does not replace those guidelines, but provides the practical security layer that sits alongside them.

ON-PREMISE OR EXTERNAL LLM: CHOOSING YOUR ARCHITECTURE

The fundamental choice every firm must make is whether to host an AI model on its own infrastructure (on-premise deployment) or to use an external, cloud-hosted service. On-premise deployment offers maximum privacy and full control over data, and is the most straightforward path to compliance with professional secrecy rules. Its drawbacks are real: significant setup cost, IT expertise requirements, and the rapid obsolescence of self-hosted models.

Using an external LLM offers access to state-of-the-art capabilities at lower infrastructure cost, but means that data leaves the firm, making contractual and regulatory compliance critical. The SAV/FSA Guidelines validate both paths, subject to strict conditions for the external option. The choice is not permanent: a well-designed infrastructure can support both simultaneously and allow the firm to evolve its approach as the technology matures.

PROFESSIONAL SECRECY: THE NON-NEGOTIABLE CONSTRAINT

For Swiss lawyers and patent attorneys, professional secrecy is not a compliance formality: it carries criminal sanctions under Art. 321 SCC (RS 311.0), and professional disciplinary obligations under Art. 13 BGFA (RS 935.61) for lawyers and Art. 10 PAG (RS 935.62) for patent attorneys. Entering client data into an AI system constitutes a disclosure to a third party, and must be treated as such. The Federal Act on Data Protection (FADP, RS 235.1) applies directly to all AI-assisted data processing.

An external provider can be treated as a legitimate auxiliary only if a set of cumulative and strict conditions are met: it must be located in Switzerland or the EU/EEA, or benefit from an adequate data transfer mechanism; prompts and responses must remain stored within those jurisdictions; the provider must be contractually bound not to share data with third parties; and it must commit to not using the firm's data for model training or fine-tuning. Where these conditions are not met, client data must not be entered - not even in a form the lawyer considers anonymised, if re-identification remains possible.

Where they are met, data minimisation remains good practice: avoid names, case numbers, and identifying details wherever the task permits it.

With the client's specific and informed consent, a lawyer may use a tool that would not otherwise meet the above criteria - but this consent must cover both professional secrecy and data protection, and cannot be assumed or inferred.

INTERNAL POLICY, GOVERNANCE AND THE SHADOW IT PROBLEM

Before any AI tool is deployed, the firm must have a written internal AI policy approved at partner level. It should specify which tools are authorised, what categories of data may and may not be used as prompts, who is authorised to use which tools, and what review process applies before AI-generated content is used or transmitted. Without such a policy, the most significant risk is not from external attackers, but from within. Collaborators who discover a useful AI tool and begin using it without authorisation, entering client data into consumer-grade systems such as free versions of commercial chatbots, represent one of the most immediate and underestimated threats to professional secrecy. This is what is known as shadow IT: the use of unsanctioned tools operating entirely outside the firm's security perimeter. The policy must explicitly prohibit this and make clear the professional consequences of non-compliance.

AI-specific risks also require dedicated training, separate from general cybersecurity awareness. Staff must understand what professional secrecy means in the context of a prompt, why verifying AI output is a professional obligation rather than a suggestion, and why deviating from the firm's approved toolset carries real consequences. Training should be periodic rather than a one-time event, given the pace at which both tools and threats evolve.

CONTRACTUAL SAFEGUARDS WITH EXTERNAL PROVIDERS

Public Terms and Conditions offer inadequate protection: they are unilaterally modifiable and rarely drafted with professional secrecy obligations in mind. A signed bilateral Data Processing Agreement (DPA) is strongly preferred. Key clauses to verify or negotiate include an explicit prohibition on using the firm's data for model training or fine-tuning, clear data residency commitments binding the provider to Switzerland or the EU/EEA, a prohibition on third-party sharing including affiliates and subcontractors, a confidentiality obligation on provider personnel, defined and short data retention and deletion timelines, and audit rights. Relevant certifications (ISO 27001 for information security management, and increasingly ISO 42001 for AI management systems) are a useful baseline indicator of a provider's maturity.

Where a DPA cannot be negotiated, the firm should at minimum document its review of the applicable T&Cs, subscribe to change notifications, and reassess its position whenever those terms are updated.

DATA MANAGEMENT: CONVERSATION HISTORY AND PROVIDER INDEPENDENCE

Another risk concerns conversation history and contextual memory. LLM performance improves as the model accumulates context about the user, their clients, and their matters, but this accumulated history is sensitive firm data. It should not be stored permanently in the provider's infrastructure. Best practice is to use a chat management layer that stores conversation history locally within the firm's own infrastructure, transmitting only the necessary context to the LLM at the time of each query, and to choose providers that commit contractually to rapid deletion of session data.

More broadly, firms should build a provider-agnostic infrastructure - one that does not create dependency on a single vendor. Tools built on open protocols such as the Model Context Protocol (MCP) allow the firm to change providers without losing its stored context or disrupting its workflows. This matters not only for flexibility but for resilience: a provider may change its pricing, its terms, or cease to operate. Vendor lock-in in a mission-critical workflow is a strategic risk that firms rarely consider until it is too late.

RESPONSIBILITY FOR AI OUTPUT: HALLUCINATIONS, LIABILITY AND INTELLECTUAL PROPERTY

The lawyer or patent attorney remains fully responsible for all work product, whether AI-assisted or not. This has two practical dimensions. The first is hallucinations: LLMs can and do produce plausible but factually incorrect legal citations, case references, or statutory provisions, typically in the same confident and authoritative tone as accurate outputs. Submitting an AI-generated brief or opinion without independent verification of every cited source exposes the lawyer to professional liability and, in litigation contexts, to sanctions from the court.

The second dimension is intellectual property. AI-generated output may incorporate content drawn from the model's training data, potentially reproducing third-party copyrighted material without attribution or licence, exposing the firm to infringement claims for content used in good faith. At the same time, under Swiss copyright law (Art. 6 URG), only a natural person can be an author: content generated by AI carries no copyright protection, and the firm cannot assert proprietary rights over purely AI-generated work product. The lawyer or patent attorney who uses AI bears both the efficiency gain and the legal risk.

PROMPT INJECTION AND EMERGING THREATS

As AI tools are integrated into legal workflows, a new category of attack becomes relevant: prompt injection. A malicious actor can embed hidden instructions within a document or email that the AI is asked to process. The model, unable to distinguish between legitimate instructions and injected commands, may be manipulated into leaking information, producing misleading summaries, or taking unintended actions. This risk is heightened as AI systems become more autonomous: a system that can autonomously send emails or access external databases on the lawyer's behalf represents a significantly larger attack surface than a passive text generator. Firms should ensure that their Information Security Expert is familiar with these AI-specific attack vectors, which are distinct from classical malware and phishing threats and require a different response.

CONCLUSION

Generative AI is not a future risk: it is a present one, already operating in most firms through approved tools, shadow IT, or both. The SAV/FSA Guidelines of 2024 provide the professional framework; this chapter provides the information security layer. The key message is straightforward: use AI, but govern it. A written policy, the right contracts, a provider-agnostic architecture, trained staff, and rigorous output verification are the minimum requirements for responsible deployment. As with all aspects of information security covered in this guide, the starting point is to work with your Information Security Expert, and, for the professional secrecy dimension, with a lawyer or patent attorney familiar with the specific obligations of the legal profession and with the IT law.