

# Governance Model: Risk Based Human Review

**Key Question: Is an elementary mistake in a material aspect tolerable?**

- No → Full human verification of the entire deliverable.
- Yes, but consequences are limited → Structured validation of key elements
- Yes, low impact → Cursory review.

| Tolerability of elementary error in material aspect | Typical examples  | Required review depth                 | Rationale  |
|---|---|---------------------------------------|--|
| Not tolerable                                       | Client advice, court or regulatory filings, board materials, high-value contracts | Full human verification               | Spot checks cannot reliably detect non-systematic errors |
| Limited tolerance (impact manageable, internal use) | Internal drafts, negotiation prep, early-stage analysis                           | Structured validation of key elements | Focused verification of material issues                  |
| Tolerable / low impact                              | Brainstorming, formatting, early idea generation                                  | Cursory review                        | Error consequences are minimal                           |

## Client Transparency Model

**Core Principles:**

- lawyers remain fully accountable for the final work product.
- full human verification is the rule for any document sent to a client, court or a regulator
- unless an exception (see below) applies: no need to disclose use of AI
- it is the lawyer's professional judgment whether an elementary AI error is legally tolerable in a certain work product or not (see governance model).
- Where the acceptability of residual risk becomes a business decision (weighing time and costs considerations against risk), that decision must be made by the client, based on transparent legal advice, and documented accordingly.
- Documentation of AI involvement serves governance, defensibility, and audit readiness. It does not constitute an admission of unreliability, but evidence of structured risk management. It must be labelled as privileged information and consider litigation and discovery aspects

## **When transparency to the client is required**

Make residual AI risk explicit if:

- A full human verification was not performed, **and**
- An elementary error in a material aspect cannot be fully excluded, **and**
- The risk is acceptable from a legal perspective and acceptance of residual risk depends on business priorities (e.g., time, cost, urgency) not legal considerations.

## **How to Make Residual Risk Explicit**

Communicate clearly and document:

### **a) Scope of AI involvement**

- What the AI was used for (e.g., drafting, summarization, clause comparison, analysis).
- Whether AI output was used as a starting draft or incorporated substantially.

### **b) Extent of human validation**

- What was fully reviewed.
- What was reviewed selectively.
- What was not independently re-verified.

### **c) Residual risk**

- Acknowledge that non-systematic elementary errors cannot be fully excluded without complete human re-verification.
- Identify which aspects would be most material if incorrect.

### **d) Worst-case consequence**

- Briefly describe the potential legal/commercial impact if a material elementary error occurred.

### **e) Alternative risk mitigation**

- What additional human review would entail.
- Estimated additional time and/or cost.
- Expected qualitative reduction of risk through full verification.

### **f) Document business risk acceptance**

- Obtain and document explicit client consent that the residual risk is accepted.
- Preserve documentation for governance and audit purposes.
- Keep in mind litigation and discovery aspects and label it privileged

Questions/feedback to [adrian.schaub@gmx.ch](mailto:adrian.schaub@gmx.ch)