

Unsafe Legal Clauses That Should Not Be Fed to AI Models

1. Vague and Ambiguous Terms

- **Clause:** "The party shall act in good faith and adhere to the best practices in all matters."
- **Why Unsafe:** The phrase "good faith" and "best practices" are **undefined** and **subject to interpretation**, making it impossible for an AI model to predict legal outcomes accurately.

2. Unfair Indemnification Clauses

- **Clause:** "The contractor shall indemnify and hold harmless the company from any and all claims, regardless of the cause."
- **Why Unsafe:** This creates **unlimited liability** for the contractor, which is unfair and legally unsound. AI models trained on such clauses may **overestimate liability risks** in contracts.

3. Excessive Penalties and Unenforceable Fines

- **Clause:** "If the employee resigns within 12 months, they shall pay a penalty equal to six months' salary."
- **Why Unsafe:** Such clauses may be considered **unconscionable** and violate labor laws, leading to **incorrect AI-based legal advice**.

4. Unreasonable Non-Compete Agreements

- **Clause:** "The employee agrees not to work in any industry-related field anywhere in the world for ten years after leaving the company."
- **Why Unsafe:** AI models trained on such clauses might **incorrectly assume all non-compete agreements are valid**, even when they are **legally unenforceable due to unreasonable restrictions**.

5. Privacy Violations

- **Clause:** "The company reserves the right to access and monitor all personal communications of the employee."
- **Why Unsafe:** This contradicts **privacy laws** such as GDPR and may lead AI models to **misinterpret legal boundaries of employee monitoring**.

6. Bias-Encouraging Risk Assessment Clauses

- **Clause:** "Decisions regarding eligibility shall be determined at the sole discretion of the company, taking into account historical risk factors."

- **Why Unsafe:** If an AI model is trained on such clauses, it may **reinforce biases** based on historical data that discriminate against specific groups.

7. Overly Broad Arbitration Clauses

- **Clause:** "All disputes, including those arising from criminal conduct, shall be settled through arbitration."
- **Why Unsafe:** Criminal matters **must be handled by courts**, not arbitration. AI trained on such clauses might **erroneously recommend arbitration for all disputes**.

8. Unclear Termination Conditions

- **Clause:** "The company may terminate the agreement if the employee does not meet expected performance."
- **Why Unsafe:** The term "expected performance" is **not defined**, leading to **subjective enforcement** that AI models cannot fairly interpret.

9. One-Sided Confidentiality Agreements

- **Clause:** "The employee shall never disclose any information about the company, but the company has the right to disclose any employee-related information as deemed necessary."
- **Why Unsafe:** This **lacks mutuality** and could lead AI to **incorrectly assess fairness in confidentiality agreements**.

10. Blanket Waivers of Liability

- **Clause:** "The customer waives all rights to sue the company under any circumstances, even in cases of gross negligence or fraud."
- **Why Unsafe:** Such clauses are **unenforceable in many jurisdictions** and may cause AI models to incorrectly assume that **waivers of liability are always valid**.

Conclusion

Feeding such **legally flawed, unfair, or ambiguous clauses** into an AI model can result in **biased legal interpretations, invalid contract generation, and incorrect legal risk assessments**. It is essential to filter out such problematic clauses before training AI models for legal applications.