Common types of bias
How they might present in an LLM

Bias	Description	Example
Algorithmic or model bias	The way the AI is trained or optimized to predict can shape how it interprets legal language.	Heavy reliance on SEC filings and public disclosures could bias outputs toward public company compliance standards, overlooking the standards of private companies.
Anchoring bias	Early outputs shape user expectations, even if flawed.	If the model's first summary is wrong, reviewers may unconsciously look for evidence to support it.
Automation bias	Users over-trust AI outputs.	Relying on AI-suggested citations without verification, assuming the system must be correct.
Confirmation bias	Models reinforce patterns seen in data rather than challenge or nuance them.	An LLM might surface only majority opinions in case law while ignoring dissenting views.
Cultural or representation bias	Some groups, topics, or perspectives are underrepresented or overrepresented in the training data.	More case law available from large U.S. courts than from smaller jurisdictions may skew outputs.
Framing bias	The way a prompt is worded influences the model's response.	Asking "Why is this contract unenforceable?" may lead the AI to assume unenforceability and search only for problems.

Legal Ops Community by Onit | Where AI meets legal innovation
Become a member