

The Case for Large Language Models in Contract Generation

S.V. Kannangara¹

Abstract

Does a contract need a mind behind it? Eliza Mik's influential working paper says yes and concludes that large language models, lacking genuine understanding, can produce only "contractual deepfakes." This article argues that Mik's technical premise is accurate but her jurisprudential conclusions do not follow from it: contract law has never required epistemically validated drafting, only functionally adequate text. Through systematic doctrinal analysis, five structural defects are identified in Mik's argument, including concessions within her own paper that are fatally inconsistent with her conclusions. In their place, this article proposes the Functional Adequacy Threshold, i.e., a context-calibrated standard grounded in the objective theory of contract and attentive to the access to justice dimension that Mik's analysis entirely ignores. The argument is not that AI contract generation is risk-free. It is that risk calls for governance, not prohibition.

Keywords: Large Language Models, Functional adequacy threshold, Objective theory of contract, Retrieval-Augmented Generation, Agentic AI, Access to justice, AI governance

Introduction

In a seminal analysis by Eliza Mik (2026), in her work of 'Contractual Deepfakes: Can Large Language Models Generate Contracts?' a serious, technically informed, and intellectually energetic paper is presented. It is also, on closer inspection, substantially less damaging to the case for AI contract generation than its rhetorical ambition suggests. The paper's technical account of Large Language Model (hereinafter referred to as "LLM") architecture is accurate, but its jurisprudential conclusions do not follow from its technical premises.

¹Assistant Lecturer – University of West London – United Kingdom / University of Sri Jayewardenepura – Sri Lanka

That gap between accurate diagnosis and overstated remedy is the subject of this article. Mik (2026) is right that LLMs lack the embodied grounding that genuine linguistic understanding, in a philosophically demanding sense, requires. She is wrong that this technical fact translates into the legal conclusion that LLM-generated contracts are unfit for purpose. Contract law has never demanded that the drafter understand the subject matter in the rich phenomenological sense that Mik's argument requires. It has demanded something considerably more modest: that the resulting text be adequate to express and govern the parties' agreement. This is also insinuated by Prof Weeramantry (1967) and agreed by other legal luminaries.

The stakes of this debate need a pragmatic grounding with real scenarios of contracts being made without key formation requirements such as 'Consensus ad idem/meeting of minds', which will be discussed in detail below. Thus, it also concerns on who has access to legally adequate contractual protection. The world's legal systems currently operate under conditions of grotesque access inequality. The World Justice Project (n.d.) estimates that an aggregate of 5.1 billion people worldwide has unmet justice needs, i.e., in practical terms, that the realistic alternative to an AI-generated contract is often no written contract at all. In the United States alone, the Legal Services Corporation (2022) has documented that low-income Americans receive no or inadequate legal help for 92 per cent of their civil legal problems. These are not background facts to be noted in passing but central facts against which any evaluation of AI contract generation must be conducted.

This article cannot and does not argue that LLMs are perfect contract drafters nor that AI-generated contracts are without risk nor even state that professional legal expertise is obsolete. It argues something more modest and more defensible: that the appropriate standard for evaluating AI contract generation is not an idealized benchmark that even experienced lawyers routinely fail to meet, but a context-sensitive threshold of functional adequacy calibrated to the realistic alternatives available to contracting parties and to the governance frameworks within which AI tools are deployed. Mik's (2026) 'veto', the argument that AI systems should not generate contracts because they lack the cognitive architecture that genuine linguistic understanding requires, fails this test.

The article proceeds as follows. Part II identifies five core structural weaknesses in Mik's argument. Part III examines what contemporary AI systems, including retrieval-augmented and agentic architectures, can actually do. Part IV proposes the Functional Adequacy Threshold framework as a positive alternative. Part V addresses the access to justice dimension that Mik's paper systematically ignores. Part VI outlines a governance response to Mik's legitimate concerns. Part VII concludes.

The Structural Defects in Mik's Argument

❖ The Philosophically Inflated Standard of Understanding

The core premise of Mik's (2026) argument is that contract drafting requires 'understanding' in a sense that LLMs cannot, as a matter of principle, possess. She grounds this claim in the 'symbol-grounding' problem, as developed by Bender and Koller (2020) in their well-known 'stochastic parrots' intervention: the argument that language models manipulate symbols without grounding them in the embodied experience of the physical world, and therefore lack genuine understanding. On the basis of this premise, Mik concludes that LLMs' contractual output lacks the meaning that a legally operative contract requires. This argument while philosophically valid has little legal relevance.

It has been trite contractual law, that it not requires, and has never required, that the drafter of a contract 'understand' the subject matter in the embodied, experiential sense that Bender and Koller (2020) identify as absent in LLMs. Lord Hoffmann's canonical account of contractual interpretation in *Investors Compensation Scheme Ltd v West Bromwich Building Society* (1997) asks what a reasonable person, knowing the background facts of the transaction, would have understood the parties to have meant. This is therefore a test of communicative adequacy not of cognitive depth. Furthermore, Weeramantry (1967) denotes that the "Contract is a promise, or promises mutually exchanged, setting up, against the promisor or promisors, duties of performance which the law will recognize or enforce at the instance or for the benefit of the promise or promises, or of a third party to be benefitted." However, he himself in reference to Roscoe Pound, Pollock, Chitty and even Treitel (Peel, 2025), denote that in creating his own definition that he took ample regard to ascertain "a due sense of the difficulties involved" that his definition too was an "attempt" and thus Weeramantry himself would agree with Mik's blunt statement of "there is no 'perfect' Contract" and thus no perfect

definition. However, the question we must ask is whether the text adequately expresses the parties' agreement in terms that a reasonable observer could understand and not whether the drafter possessed a philosophically verified form of linguistic understanding.

Applied uniformly, Mik's understanding requirement would invalidate the drafting work of a substantial proportion of practicing lawyers. To exemplify, consider the solicitor two years out of law school who drafts a charterparty without ever having been to Sea or the corporate associate who drafts a pharmaceutical licensing agreement without any training in biochemistry or the tax lawyer who drafts a complex structured finance instrument without having managed a balance sheet. Even where the contract is formally valid, the subjective realities of the parties may be entirely divorced from its terms. Consider a university academic, A, whose understanding of their institutional relationship is framed entirely in terms of pedagogical duty, while the university's understanding of the same instrument is expressed through quantifiable performance metrics. Neither party has misrepresented their intention; both have appended their signatures. The Law of Contract does not require a meeting of souls, merely a meeting of signs. Reverting to the maxim, 'all contracts are agreements but not all agreements are contracts', this maxim essentially acts as the gatekeeper of legal validity. Thus, an agreement generated by AI remains a mere collection of sophisticated syntax until the objective act of signing occurs. But if the AI-generated contract is divorced from the subjective realities of the signatories, as with A, then we must ask whether the law is facilitating a meeting of minds or a 'collision of signatures'.

The doctrine of consensus ad idem has increasingly transitioned from a psychological requirement to a functional fiction. As demonstrated by A's 'unqualified assent' to a contract he did not read, the 'grundnorm' of Contract Law; the 'Objective Theory of Contract' operates on an inference of intent. It prioritizes the "outward manifestation" of agreement over the internal 'preconception' of the parties (Yonjan, 2019). This concept can be easily replicated in the realm of AI generated contracts, as can be seen by the operability of multiple firms not just using AI to draft contracts but also have AI firms who draft such Electronic or Smart contracts (Cuccuru, 2017; Pollins, 2026). Coming back to the foundational issues, none of these practitioners understands the subject matter of their contracts in the embodied experiential sense that Mik's argument treats as necessary. They operate, as lawyers invariably do, by ironically, matching patterns against precedent,

applying legal categories to facts they understand only in legal terms and relying on client instructions to supply the commercial understanding they themselves lack. A statistical analysis in 2025 (Ambrogio, 2025; Legal AI Benchmarking, n.d.), not only proves this fact but strengthens this position via the synergy of platform workflow support, output usefulness and output reliability becomes "max utility", it is this concept that is the genesis of AI tool utility. Thus, as a tool itself it is being used more than 97% of lawyers (American Bar Association, 2025) and it even outperforms a human made contractual first draft by 3% and top AI models consistently outperform the average lawyer (Ambrogio, 2025; Legal AI Benchmarking, n.d.).

The aforementioned study disarms Mik's core argument in a statistical-quantitative manner, however, it is noteworthy that the version that might withstand this critique is not that LLMs lack 'understanding' in the abstract, but that they lack the specific functional competences that adequate contract drafting requires: knowledge of applicable law, awareness of the commercial context and sensitivity to the consequences of particular formulations. However, while these are valid concerns about functional adequacy, it is not about philosophical architecture. For the pragmatist, the functional adequacy is, as we shall argue in Part IV, precisely the right test.

❖ **The Rigged Comparison: Templates and the Privileged Baseline**

Mik's (2026) efficiency argument turns on a comparison between AI contract generation and the 'customization of templates.' She argues that templates 'used in law firms often represent years of drafting experience' and have been refined through successive commercial transactions, whereas AI-generated documents lack this accumulated wisdom.

The comparison is constructed against a baseline that systematically excludes the population for which AI contract generation is most valuable. Mik is comparing AI generation against the gold standard of professional legal practice, bespoke template customization by experienced solicitors with institutional knowledge of the relevant transaction type. This resource is simply unavailable to the vast majority of contracting parties. Legal Aid Agency data indicates that fewer than 15 per cent of civil legal matters in England and Wales receive professional assistance (Legal Aid Statistics England and Wales Bulletin, 2024). The Legal Services Corporation's 2022 Justice Gap study found that low-income Americans received no

or inadequate legal help for 92 per cent of their civil legal problems (Legal Services Corporation, 2022).

The methodological implication is significant. Mik's comparison is structurally equivalent to evaluating over-the-counter analgesics by comparing them to treatment in a specialist pain clinic, while ignoring that the specialist clinic is unavailable to most of the population. The relevant comparison, for most users of AI contract generation tools, is not "AI generation versus law firm template" but "AI generation versus no written contract at all" or "AI generation versus a template of entirely unknown provenance downloaded from the internet." Even within the complex commercial contracting context that Mik assumes, the template comparison is less straightforward than she suggests. The identification of the appropriate template for a novel transaction type requires legal judgment that template repositories do not supply automatically. Adapting a standard template to the idiosyncrasies of a particular commercial relationship, parties in different jurisdictions, unusual risk allocation requirements, novel payment structures requires exactly the kind of expertise that Mik assumes templates make available.

Mik herself gestures toward this problem when she acknowledges that 'many parties do not contemplate the possibility of litigation and sign arbitrary documents downloaded from the Internet.' But having observed that reality, she offers no analysis of what it implies for the normative question she is addressing. The tacit assumption, that the relevant comparator for AI contract generation is expert legal advice rather than the informal arrangements and downloaded templates that most contracting parties actually use, actually does a great deal of work in Mik's argument that it never earns. Yet, all of these arguments get defenestrated by the 2025 paper showing that not only many lawyers use AI for contract generations and AI has now become better at it than the average lawyer, marginally though it may be (Ambrogio, 2025; Legal AI Benchmarking, n.d.).

❖ **The Lethal Concessions: Mik's Own Admissions Are Structurally Fatal**

The most decisive rebuttal to Mik's argument comes from Mik's own paper. In a sequence of admissions that she never reconciles with her overall thesis, Mik (2026) concedes: that LLMs can generate documents that 'look like contracts' and 'may be adopted by parties'; that once adopted, parties are bound because 'it is irrelevant whether provisions were generated by LLM or drafted by lawyer'; that 'the objective theory of contract disregards technical provenance of text'; that an LLM-generated

document 'could contain minimum substance required' and be 'prima facie enforceable'; and that 'with unlimited resources... LLMs could generate viable contracts, not just contractual deepfakes.'

These are not minor qualifications at the margins of Mik's argument. They are concessions that when taken together, establish that LLMs can generate enforceable contracts and that the law treats such contracts identically to contracts of any other provenance, and that viability is, in Mik's own analysis, a function of the resources and expertise applied to review and not a consequence of the architectural impossibility she elsewhere asserts.

The structural logic of Mik's paper, once these concessions are pressed, turns out to be this: LLMs can generate enforceable contracts; the law is indifferent to their technical provenance; but they should not be used because the resulting contracts may be commercially inadequate. This is a considerably less interesting argument than the one Mik appears to be making. It is, in substance, an argument for quality regulation and informed usage guidance. It is not an argument for the sweeping condemnation that Mik's paper advertises.

Thus, despite her evident intentions, her paper actually establishes the legal legitimacy of AI contract generation while identifying the conditions under which supplementary governance is required. The paper's architecture works against its own conclusions.

❖ Dismissing Retrieval Augmented Generation and Agentic Architectures

Mik (2026) addresses what she calls 'potential solutions' on fine-tuning on legal texts, sophisticated prompting and retrieval-augmented generation (RAG) in a very brief exposition. Her dismissal confidently states that RAG systems 'cannot distinguish between relevant and irrelevant information' and cannot 'apply' retrieved legal knowledge. For support, she relies primarily on a single 2025 benchmarking study.

More fundamentally, Mik's analysis of 'potential solutions' is conducted entirely within the framework of single-pass LLM text generation: the model receives a prompt and produces output. This framework does not describe the AI systems that legal technology companies are actually deploying. Contemporary AI-assisted

contract drafting systems operate via different architectures. In an usual agentic legal drafting system, a structured intake process specifies transaction parameters; an orchestrating AI agent coordinates multiple specialised tool calls; the system retrieves jurisdiction-specific statutory materials, precedent documents, and applicable case law from curated legal databases (Lewis et al., 2020), afterwards, an iterative self-correction loops check generated provisions for internal consistency, following which, citation verification confirms that retrieved legal sources are accurate and finally structured human review checkpoints integrate lawyer judgment at defined stages (Schwarcz et al., 2025).

The distinction between a single-pass LLM and a multi-agent agentic system with human-in-the-loop review is vital as the distinction between a junior associate drafting unaided and a supervised drafting process with senior partner review and precedent library access. Mik's analysis of the former does not speak to the latter.

Schwarcz and others (2025) performed a randomised controlled trial in 2025, the results of which are remarkable to say the least, which Mik herself cites is much more instructive and prudent than her brief treatment allows. The study, the first of its kind (in legal settings), found that both reasoning-model AI tools (OpenAI's o1-preview) and RAG-powered legal AI tools (Vincent AI) significantly enhanced legal work quality across multiple task types, with o1-preview yielding quality improvements of 10-28% across four assignments and RAG tools providing improvements of 8-15%. Productivity gains were substantial, ranging from 50 to 110%. Critically, the study found that quality improvements from reasoning models were concentrated in litigation-oriented tasks rather than transactional work, this finding itself identifies where further development is needed, rather than establishing impossibility.

The benchmarking literature more broadly supports a picture of rapid development rather than architectural impossibility. Katz and others (2024) denoted that GPT-4 substantially outperforms prior models on the Uniform Bar Examination, a test of structured legal reasoning and analysis, shows LLMs developing functional legal reasoning capacities.

Of course, it is an entirely another argument whether we are to re-evaluate the learning parameters and whether knowledge and application of said knowledge is used in the bar exam, or simply a regurgitation of the information. The latter

argument would probably be Mik's stance. The critique to this probable assertion would follow the deductive reasoning that if the bar exams which are supposed to evaluate a candidate's aptitude, where the students perform consistently lower than the LLM, and then as novice attorneys/lawyers, make contracts that are also consistently of a lower standard. Zheng and others (2025) showed reasoning-focused legal retrieval benchmark demonstrates steady improvement in LLMs' ability to identify relevant legal materials even when lexical similarity between query and document is low. Jonathan Choi's (2024) empirical work on legal text clarity which elucidated that legal language is frequently indeterminable even for the experts, judges, essentially cuts against the assumption that the adequacy of contract drafting can be assessed against the kind of determinate standard Mik appears to presuppose.

But none of this establishes that AI contract generation has solved the problems that Mik posits. However, it does establish that Mik's conclusions were drawn on the basis of a partial and rapidly dating technological assessment.

❖ **The Category Error: Evaluating Generation Against an Impossible Ideal**

A prevalent structural defect in Mik's (2026) argument is the implicit comparison of AI-generated contracts to an idealised standard of human drafting that does not, in fact, exist at scale. Mik's 'viable contract'...one that 'fits' the transaction and applicable law, reconciles the parties' conflicting interests and addresses the consequences of non-performance, this then becomes a formidably demanding specification. While it is the contract that every client imagines their lawyer is producing, it is not the contract that most lawyers actually produce.

The courts are satiated with disputes arising from contracts drafted by qualified solicitors that fail adequately to address the parties' intentions, contain ambiguities generating years of costly litigation or produce commercially nonsensical results when read literally. The very principle of purposive interpretation, as articulated in *Rainy Sky SA v Kookmin Bank* (2011), is premised on the observation that literal contractual text frequently fails to reflect commercial common sense, that lawyer-drafted contracts, in other words, regularly fail Mik's viability standard. Mik herself cites *Wood v Capita Insurance Services Ltd* (2017) for the proposition that even contracts drafted by lawyers may be imperfect without appearing to register what this concession implies for the standard she is applying to AI.

Thus, the appropriate comparison is not AI generation against the ideal contract but AI generation against the realistic distribution of contracts actually produced in the relevant context. For the sophisticated commercial context that Mik assumes, the comparison is AI-assisted drafting against lawyer-reviewed drafting. For the "access to justice" context that Mik ignores, it is AI generation against no written contract at all. Mik's category error has the structure of a common rhetorical tactic, which is to identify the limitations of a new technology by measuring it against a standard that the technologies it displaces or that are unavailable to its principal users, which do not themselves meet. This is not an intellectually serious evaluative framework.

The Positive Case: Functional Competence in AI Contract Systems

❖ What Agentic AI Systems Can Actually Do

The AI systems that Mik (2026) analyses are essentially base large language models that generate text in response to a single natural language prompt. This technology, as it existed in 2022 and 2023, warranted considerable caution. Mik's assessment of those early systems is not unreasonable. The difficulty is that this is no longer a complete or accurate description of the AI systems being deployed in legal practice.

Novel AI-assisted contract drafting systems deploy a fundamentally different architecture. A structured intake process captures the key parameters of the transaction, where the parties, jurisdiction, transaction type, commercial context, specific risk allocation requirements. An AI orchestrator coordinates multiple specialised calls: to legal databases for jurisdiction-specific statutory materials; to precedent repositories for transaction-type-specific clause libraries; to compliance checking tools for mandatory legislative requirements; to citation verification systems to confirm that retrieved legal sources are current (Lewis et al., 2020). Iterative self-correction loops identify and flag provisions that are internally inconsistent or that produce commercially anomalous outcomes when read in combination and structured human review checkpoints integrate lawyer judgment at defined stages, ensuring that AI-generated text is reviewed before adoption rather than accepted uncritically (Schwarcz et al., 2025).

This process is not, to borrow Mik's vivid example of a system operating by instinct in a medium it cannot conceptually apprehend, the octopus in the water. It is closer to a sophisticated document assembly system with AI-driven customisation,

integrated legal validation and mandatory human oversight at defined intervention points.

The question of whether such a system can generate 'viable contracts' in Mik's sense is not a question about whether base LLMs understand delivery clauses. It is a question about whether the end-to-end system produces a viable output that is adequate for its purpose. These are different questions, and Mik addresses only the first.

❖ **Empirical Evidence on AI-Assisted Legal Drafting**

The empirical picture on AI-assisted legal drafting is more nuanced and considerably more encouraging, than Mik's (2026) paper suggests.

Schwarcz and others' (2025) study with its novel approach of randomised controlled trials, found that both RAG-powered and reasoning-model AI tools significantly enhanced the quality of legal work across multiple task categories, a result that the authors describe as a 'marked contrast with previous research examining older large language models.' The AI tools boosted participants' productivity in five out of six legal tasks with particularly strong effects in drafting and analytical tasks. The finding that quality improvements were concentrated in litigation-oriented matters rather than transactional tasks identifies a specific limitation that further development can address rather than an architectural ceiling. Furthermore, the key quantitative analysis backed paper in 2025 (Ambrogio, 2025; Legal AI Benchmarking, n.d.), proves that AI contracts are more than a gimmick but effectively beating the average lawyer, consistently. Introduce a paid version of a AI tool, and this gap between the AI tool and the average lawyer widens significantly.

Katz and others' (2024) demonstration that GPT-4 substantially outperforms prior models on the Uniform Bar Examination, beating human test-takers in five of seven subject areas on the multiple-choice component and achieving a 4.2/6.0 average on the open-ended essay components, shows LLMs developing functional capacity for the kind of structured legal reasoning that contract drafting requires. The aforementioned study, which documented GPT-3.5's performance on the multiple-choice bar exam, established the baseline from which this improvement has been measured.

The hallucination literature, which Mik correctly cites as a concern (Dahl et al., 2024), provides important context for responsible deployment rather than grounds for categorical rejection. In a recent study, the documentation of LLM hallucination rates of 58% or higher in open-domain legal queries is a serious finding but it concerns factual claims about case law and not the generation of contractual text from structured inputs. The distinction is vital: a contract drafting tool that receives specific transaction parameters, retrieves jurisdiction-appropriate clause libraries and generates text conforming to retrieved templates faces a different hallucination risk assessment than a chatbot answering open-ended questions about case law.

None of this evidence establishes that AI contract generation is superior to expert human drafting for complex, high-value, bespoke commercial transactions. It does establish that the categorical conclusion that AI 'should not' be used for contract generation is not supported by the empirical literature. What the evidence supports, instead, is context-sensitive deployment, calibrated oversight, and ongoing quality monitoring: the elements of a governance framework, not the grounds for a prohibition.

❖ The Functional Grounding of Legal Language

Mik's (2026) symbol grounding argument derived from Bender and Koller (2020), identifies a genuine philosophical problem. LLMs do not ground language in embodied experience of the physical world. That is correct. But its legal implications are more limited than Mik suggests.

Consider the concept of 'delivery' that Mik uses as her central example of a concept LLMs cannot understand without embodied experience. Mik argues that understanding delivery requires prior knowledge of physical objects being transported from one place to another. This is accurate as a philosophical account of how most humans acquire the concept. But the legal concept of 'delivery' is not the everyday concept. It is a technical legal concept defined by statute, such as the Sale of Goods Act 1979, elaborated by decades of case law and thus operationalised through standard contractual formulations that are extensively documented in legal sources. A system that has internalised the word distributions of legal texts about delivery via statutes, cases, standard form contracts, practitioner guides, possesses a functional representation of the legal concept of delivery that is sufficient for the purpose of generating legally appropriate delivery clauses.

Whether this constitutes 'understanding' in the philosophically demanding sense that Koller and Bender (2020) have in mind is, to put it gently, not the question that contract law asks. Whether it is adequate for the purpose of generating a workable delivery clause in a standard sale of goods contract is an empirical question. The empirical evidence (Schwarcz et al., 2025; Katz et al., 2024) suggests it frequently is.

The Functional Adequacy Threshold: A Positive Framework

❖ Principles of the Framework

The Functional Adequacy Threshold (hereinafter denoted as 'FAT') framework proposes that the appropriate standard for evaluating AI-generated contracts is not whether they meet an ideal of perfect transactional sufficiency, i.e., a standard that human-drafted contracts routinely fail but whether they meet a context-specific threshold of functional adequacy calibrated to transaction value and complexity, applicable legal framework, the degree of human oversight needed and applied to AI output (as some firms already do) and the realistic alternatives available to contracting parties.

At the highest level of the framework, a complex, high-value, rights-intensive commercial transactions in sophisticated jurisdictions, the FAT is demanding. AI-generated text in this context should be treated as input to a supervised drafting process, not as a finished product. Mandatory human review by a qualified legal professional is a condition of responsible deployment. This position is broadly consistent with Mik's (2026) own, and we do not contest it. At intermediate levels, the standard commercial transactions of moderate value in established transaction types, with access to legal review where AI-assisted drafting within a structured human-oversight architecture satisfies the FAT. Something which most AI law firms who make such contracts do (Daniels, n.d.; Pollins, 2026). An AI system generates an initial draft calibrated to transaction type and jurisdiction; a human reviewer applies contextual judgment to adapt the draft to the specific commercial context, then the result is a document that is functionally adequate for the transaction's legal purposes. Thus, the empirical evidence on AI-assisted legal drafting supports cautious optimism for this tier (Schwarcz et al., 2025).

At the base of the framework, low-value or informal transactions, access contexts in which professional legal advice is unavailable and standard-form consumer contracts for which the relevant question is whether the consumer is bound by

terms they have received notice of (*L'Estrange v F Graucob Ltd*, 1934), then the FAT is calibrated accordingly. The relevant comparison is AI generation against no written contract, a contract of unknown provenance downloaded from the internet even a jurisdictionally inappropriate standard form. An AI-generated document that produces an enforceable written record of the transaction's commercial elements, in plain language, accessible at nominal cost, satisfies the FAT for this tier even if it does not achieve the transactional precision of a bespoke lawyer-reviewed contract.

❖ The Objective Theory of Contract as Foundation

The FAT framework is anchored in a principle that Mik (2026) herself articulates but fails to develop: the "objective theory" of contract. Under this principle, which governs contract formation and interpretation in English law, the courts determine contract content by reference to what a reasonable person in the position of the parties would have understood the contract to mean, not by reference to the subjective mental states of the parties or the cognitive processes of the drafter. This principle is fundamental to Lord Hoffmann's interpretation canon in *Investors Compensation Scheme* (1997) and is applied consistently through *Rainy Sky* (2011) and *Wood v Capita* (2017).

The technical provenance of contractual text, i.e., whether it was typed by a lawyer, dictated, produced by autocomplete or generated by an LLM, is irrelevant to the objective theory. As Mik herself concedes, once parties adopt an LLM-generated document as their contract, 'it is irrelevant whether provisions were generated by LLM or drafted by lawyer' the law enforces it as written (Mik, 2026). This is not merely a technicality. It reflects a deep structural commitment of common law contract doctrine to functional adequacy over formal pedigree: even Weeramantry (1967) states that the law enforces contracts because they represent the parties' outward manifestation of agreement, not because they were produced through epistemically validated cognitive processes.

Mik's argument, at its core, is that AI systems should not be used for contract generation because their outputs may be inadequate. The objective theory of contract, which Mik endorses, is the very principle that makes her concern a governance question rather than a legal impossibility question. AI-generated contracts can be legally binding and they are but the question is whether they should be, under what conditions and finally but most importantly, with what safeguards. That is a governance question for which a governance answer is appropriate.

The Dimension Mik Ignores: Access to Justice and The Global Contracting Gap

❖ The Global Scale of Legal Exclusion

The most significant deficiency in Mik's (2026) analysis is distributional: it is the complete absence of any engagement with the access to justice dimension of AI contract generation. Mik's paper proceeds as if the relevant question is whether AI can produce contracts of the quality that sophisticated commercial actors expect from their legal advisers. It does not ask whether AI can produce contracts of adequate quality for the vast majority of the world's commercially active population, who have no realistic access to legal advisers at all.

The World Justice Project (n.d.) estimates that 5.1 billion people worldwide have unmet justice needs, this is anything but a marginal figure, but the majority of the world's adults. For these 5.1 billion, the realistic alternative to an AI-generated contract is not a contract reviewed by senior counsel. It is an informal agreement with unrecorded, disputed, and frequently unenforceable terms; a handwritten document of uncertain legal effect. A form downloaded from an unreliable website that is appropriate to a different jurisdiction and a different transaction type. The Legal Services Corporation's (2022) data on low-income legal needs in the United States, one of the world's wealthiest and best-resourced legal systems, gives some indication of the scale: 92% of civil legal problems faced by low-income Americans go unaddressed.

In this context, an AI system that generates an enforceable written record of a commercial transaction, in plain language, calibrated to the relevant jurisdiction, accessible at nominal cost or no cost, does not generate a "contractual deepfake". It generates something practically transformative for the world's commercially active population that lacks access to legal services: a written contract.

B. The Paternalism Problem

Mik's (2026) argument has a structural paternalism problem that is worth identifying explicitly. The argument is, that AI-generated contracts should not be used because they may be of lower quality than contracts produced by expert human counsel. Applied consistently, this reasoning would prohibit: contracts drafted by non-lawyers; standard form contracts used without legal advice; contracts drafted by junior lawyers without adequate supervision and, in truth, the vast majority of

contracts that have governed commercial relationships throughout recorded history.

Law has generally declined to invalidate such contracts, for an obvious reason: the alternative, which is a legal system in which contractual protection is available only to those who can afford expert legal advice would reproduce and entrench access inequalities rather than address them. Mik's argument, whatever its intentions, has exactly this structure. The paternalism is particularly acute when directed at AI deployment in developing economies.

Consider a small business owner in Jaffna with no access to legal counsel, not because they have chosen to forego it, but because it is simply not available at any price they can pay. Mik argues that this person should not use an AI tool to generate a written record of a commercial transaction because the AI-generated document may be imperfect. The alternative she implicitly offers is a professional legal counsel, bespoke template customisation by experienced lawyers, which is not available to this person. The AI tool is available. Whatever its imperfections, it is substantially better than nothing.

Mik has no response to this, because she has not engaged with the question. The gap in her analysis is not a minor oversight. It is the central question that a normative evaluation of AI contract generation must address.

A Governance Response to Legitimate Concerns

None of the arguments advanced in this article should be taken to imply that AI contract generation is without risk or that Mik's (2026) concerns about AI-generated contractual inadequacy are unfounded. They are fecund observations that are founded, thus we cannot defenestrate the questions or observations raised, however, as mentioned prior, her own article kills her central theme. The hallucination literature documents real failures (Dahl et al., 2024). The empirical evidence on AI performance in transactional tasks is more limited than in litigation tasks (Schwarcz et al., 2025). These concerns call for governance responses, not for prohibition.

❖ Disclosure Requirements

Parties to contracts generated with AI assistance should be required to disclose the AI-assisted provenance of the document to their counterparties. This requirement serves multiple functions: Firstly, it enables counterparties to apply appropriate scrutiny. Secondly, it creates incentives for generating parties to take review seriously and finally it builds the empirical evidence base needed to assess AI-generated contract performance over time. Disclosure requirements apply regardless of the degree of human review subsequently applied.

❖ Tiered Oversight Requirements

Consistent with the FAT framework proposed in Part IV, oversight requirements should be calibrated to the value and rights-sensitivity of the transaction. For transactions above specified value thresholds, AI-generated contracts should require review and approval by a qualified legal professional before adoption. For consumer contracts and transactions governed by mandatory legislative requirements, such as the Consumer Rights Act 2015, the Unfair Contract Terms Act 1977 and their equivalents in other jurisdictions (such as Consumer Affairs Authority Act No. 9 of 2003 of Sri Lanka), AI generation should be subject to automated compliance checking against jurisdiction-specific rule-sets as a condition of deployment. For lower-value, informal, and access-context transactions, the oversight requirement should be informational rather than structural: parties should receive plain-language information about the limitations and risks of AI-generated documents, but should not be prohibited from adopting them.

❖ Professional Responsibility Rules

Legal professionals who deploy AI tools in drafting work retain their full professional responsibility obligations to clients. A solicitor who uses an AI system to generate a contractual first draft and presents it to a client without adequate review has not reduced their professional responsibility by virtue of the AI's involvement. AI is a tool and thus the professional's judgment remains the operative standard of care. Professional regulators in England and Wales, Hong Kong, and Australia are developing guidance in this area (Premalal & Widanagama, 2025). The development of coherent cross-jurisdictional standards for AI use in legal practice is an urgent professional responsibility priority.

❖ The EU Artificial Intelligence Act and Regulatory Implications

The governance framework proposed above does not operate in a legislative vacuum. The EU Artificial Intelligence Act (Regulation (EU) 2024/1689), which entered into force on 1 August 2024 with phased application through 2026, provides partial legislative instantiation of the tiered oversight and disclosure obligations advanced in this Part. The Act's risk-tiered architecture: distinguishing prohibited practices, high-risk systems, general purpose AI models, and minimal-risk applications mirrors the FAT framework's calibration of oversight requirements to transaction value and complexity. Article 50 of the Act imposes transparency obligations on providers of AI systems that interact with natural persons, requiring clear disclosure of AI involvement where a reasonable person might otherwise assume they are dealing with a human. These obligations, applied to AI-assisted contract generation in a consumer context, give existing legislative expression to the disclosure requirements proposed in Part VI.A.

The Act does not presently classify AI contract drafting tools as high-risk systems under Annex III, which reserves that designation for AI deployed in the administration of justice and AI systems used in employment, credit, and essential services access decisions. This classification may warrant reconsideration as AI contract generation scales into consumer and micro-enterprise contexts, in which contracting parties lack the sophistication to evaluate AI-generated terms or to identify when those terms inadequately reflect their commercial intentions. The governance framework proposed in this article, i.e., calibrated disclosure, tiered oversight, professional responsibility obligations, provides the conceptual architecture for regulatory bodies and legal professional bodies developing sector-specific guidance under the Act's framework provisions. The alignment between the FAT framework and the Act's regulatory logic is not coincidental: both proceed from the foundational premise that the appropriate response to risk generated by AI systems is proportionate, context-sensitive governance, not categorical prohibition.

Conclusion: Functional Adequacy – The Pushback

Eliza Mik (2026) has written a genuinely valuable paper. Her technical analysis of LLM architecture is accurate. Her identification of the risks of AI-generated contractual inadequacy is substantively well-grounded. Her call for scepticism about uncritical AI deployment in legal contexts is timely and reasonable. The

difficulty is that none of these contributions requires the sweeping conclusion she draws from them and several of her own admissions are inconsistent with it.

The argument that AI systems cannot generate viable contracts because they lack the embodied grounding that genuine linguistic understanding requires, is a stretch and proves too much. Applied consistently, it would equally invalidate the drafting work of non-specialist lawyers, lay parties, and the vast majority of contracting parties throughout commercial history. Contract law has never required that contractual text be produced through epistemically validated cognitive processes. It has required that text be adequate to express the parties' agreement and govern their transaction.

Whether AI-generated text meets that functional standard is an empirical question, not an architectural impossibility. The empirical evidence does not support Mik's sweeping pessimism. It supports a nuanced picture: AI tools can significantly enhance legal work quality and efficiency (Schwarcz et al., 2025); their performance is uneven across task types; RAG and agentic architectures substantially address the limitations of single-pass LLMs (Lewis et al., 2020); and rapid development continues (Katz et al., 2024; Zheng et al., 2025). These findings do not establish that AI contract generation is ready for all contexts without safeguards. They establish that the appropriate response is a governance framework, not a prohibition.

The FAT framework proposed here provides that governance response. It calibrates oversight requirements to the value and complexity of the transaction, mandates disclosure of AI provenance, preserves professional responsibility obligations in full, and most vitally, takes seriously the access to justice dimension that Mik's paper ignores. For the 5.1 billion people who lack meaningful access to justice globally (World Justice Project, n.d.), an imperfect AI-generated contract is not a 'deepfake.' It is, literally and practically, better than nothing. The existence of AI law firms using such contracts itself show that the assertion is false in reality itself (Daniels, n.d.; Pollins, 2026).

The debate about AI in legal practice will not be resolved by arguments from philosophical principle about the impossibility of machine understanding. It will be resolved by empirical evidence about what AI systems actually do, governance frameworks that take risk seriously without foreclosing access and an honest reckoning with the distributional consequences of legal exclusion at global scale.

Mik's paper contributes more than she intends: by establishing the legal legitimacy of AI contract generation, acknowledging the capacity of LLMs to produce enforceable documents and identifying the conditions under which risks can be managed, she has provided the foundations for the governance framework that her paper was designed to reject.

References

- Ambrogio, B. (2025, September 18). AI tools match or exceed human lawyers in contract drafting benchmark study. LawSites. <https://www.lawnext.com/2025/09/ai-tools-match-or-exceed-human-lawyers-in-contract-drafting-benchmark-study.html>
- American Bar Association. (2025). How lawyers are using AI to draft better contracts faster. https://www.americanbar.org/groups/law_practice/resources/law-practice-today/2025/october-2025/lawyers-using-ai-to-draft-contracts/
- Austen-Baker, R. L. P. S. J. (2002, June 1). Gilmore and the strange case of the failure of contract to die after all. Social Science Research Network. <https://papers.ssrn.com/abstract=1424476>
- Bender, E. M., & Koller, A. (2020). Climbing towards NLU: On meaning, form, and understanding in the age of data. In Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics (pp. 5185–5198). Association for Computational Linguistics. <https://aclanthology.org/2020.acl-main.463/>
- Black, N. (2025). Access to justice 2.0: How AI-powered software can bridge the gap. American Bar Association. <https://www.americanbar.org/groups/journal/articles/2025/access-to-justice-how-ai-powered-software-can-bridge-the-gap/>
- Choi, J. (2024). Measuring clarity in legal text. University of Chicago Law Review, 91(1). <https://chicagounbound.uchicago.edu/uclrev/vol91/iss1/1>
- Consumer Affairs Authority Act No. 9 of 2003 (Sri Lanka).
- Consumer Rights Act 2015 (United Kingdom).
- Cuccuru, P. (2017). Beyond Bitcoin: An early overview on smart contracts. International Journal of Law and Information Technology, 25(3), 179–195. <https://doi.org/10.1093/ijlit/eax003>
- Dahl, M., Magesh, V., Suzgun, M., & Ho, D. E. (2024). Large legal fictions: Profiling legal hallucinations in large language models. Journal of Legal Analysis, 16(1), 64–93. <https://doi.org/10.1093/jla/laee003>
- Daniels, R. (n.d.). Introducing Crosby, the world's first hybrid law firm. Crosby. <https://crosby.ai/post/introducing-crosby-the-worlds-first-hybrid-law-firm>
- European Union. (2024). Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence (Artificial Intelligence Act). Official Journal of the European Union, L 1689.
- Investors Compensation Scheme Ltd v West Bromwich Building Society [1997] UKHL 28, [1998] 1 All ER 98, [1998] 1 WLR 896.
- Katz, D. M., Bommarito, M. J., Gao, S., & Arredondo, P. (2024). GPT-4 passes the bar exam. Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences, 382(2270), 20230254. <https://doi.org/10.1098/rsta.2023.0254>
- Legal Aid Statistics England and Wales Bulletin. (2024). Legal aid statistics England and Wales bulletin Jan to Mar 2024. GOV.UK. <https://www.gov.uk/government/statistics/legal-aid-statistics-quarterly-january-to-march-2024/legal-aid-statistics-england-and-wales-bulletin-jan-to-mar-2024>
- Legal AI Benchmarking. (n.d.). Benchmarking humans & AI in contract drafting. <https://www.legalbenchmarks.ai/research/phase-2-research>
- Legal Services Corporation. (2022). The unmet civil legal needs of low-income Americans: The justice gap report. <https://justicegap.lsc.gov/resource/2022-justice-gap-report/>
- L'Estrange v F Graucob Ltd [1934] 2 KB 394.

Lewis, P., Perez, E., Piktus, A., Petroni, F., Karpukhin, V., Goyal, N., Küttler, H., Lewis, M., Yih, W., Rocktäschel, T., Riedel, S., & Kiela, D. (2020). Retrieval-augmented generation for knowledge-intensive NLP tasks. In Proceedings of the 34th International Conference on Neural Information Processing Systems (pp. 9459–9474). Curran Associates Inc. <https://dl.acm.org/doi/10.5555/3495724.3496517>

Mik, E. (2026). Contractual deepfakes: Can large language models generate contracts? arXiv. <https://doi.org/10.48550/arXiv.2602.09384>

Peel, E. (2025). Treitel, the law of contract (16th ed.). Sweet & Maxwell.

Pollins, M. (2026). *10 AI law firms to watch in 2026*. Lupl. <https://www.lupl.com/blog/10-ai-law-firms-to-watch-in-2026/>

Premalal, N., & Widanagama, B. (2025). Artificial intelligence, the legal profession and access. Peradeniya Law Journal, 1, 9.

Rainy Sky SA & Orsd v Kookmin Bank [2011] UKSC 50.

Sale of Goods Act 1979 (United Kingdom).

Schwarcz, D., Choi, J. H., & Ouyang, A. (2025, March 2). AI-powered lawyering: AI reasoning models, retrieval augmented generation, and the future of legal practice. Social Science Research Network. <https://doi.org/10.2139/ssrn.5162111>

Unfair Contract Terms Act 1977 (United Kingdom).

Weeramantry, C. G. (1967). The law of contracts (Vol. 1). Lawman (India).

Wood v Capita Insurance Services Ltd [2017] UKSC 24.

World Justice Project. (n.d.). Measuring the justice gap. <https://worldjusticeproject.org/our-work/research-and-data/measuring-justice-gap>

Yonjan, Y. K. (2019, August 14). An analysis on major elements of a valid contract under Muluki Civil Code, 2074. Social Science Research Network. <https://doi.org/10.2139/ssrn.3437233>

Zheng, L., Guha, N., Anderson, J., Henderson, P., & Ho, D. E. (2025). A reasoning-focused legal retrieval benchmark. In Proceedings of the Symposium on Computer Science and Law (pp. 1–12). Association for Computing Machinery. <https://doi.org/10.1145/3709025.3712219>

Article Information:

Received	21-Jun-2026
Revised	27-Aug-2026
Accepted	6-Sept-2026
Published	30-Sep-2026

Declarations:

Author's Contribution:

- **Conceptualization, and intellectual revisions. Data collection, interpretation, and drafting of manuscript**
- The author agrees to take responsibility for every facet of the work, making sure that any concerns about its integrity or veracity are thoroughly examined and addressed

▪ **Conflict of Interest:** NIL

▪ **Funding Sources:** NIL

Correspondence:

S.V. Kannangara

suweenvinsuka@sjp.ac.lk
