

GenAI and commercial contracting: what's the point?

Alex Hamilton, CEO Radiant Law

November 2023

"When the facts change, I change my mind. What do you do, sir?"

John Maynard Keynes¹

Generative AI² is all the rage. OpenAI stormed the legaltech world when it released GPT-4³ in March 2023 and a set piece ensued: prophets breathlessly claimed that lawyers were about to be replaced by robots; moderates cautioned that it would just be the lawyers who didn't get with the program(me); lawyers pointed to ChatGPT inventing cases and went back to filling in their timesheets; futurists claimed that the next version would solve everything; pragmatists pointed out that lawyers can't even use styles in Word. So business as usual, really, for discussions in the legaltech world.

Personally, I was shocked.

¹ Or probably Paul Samuelson in this form, [who still attributed it to Keynes](#).

² Generative AI (or GenAI) is the hot new domain within AI where trained models (using massive amounts of example data) are able to generate new text, images, audio, video etc. The terminology gets complicated: ChatGPT is a chat front-end for a text-based example of GenAI, underpinned by different large language models (LLMs), for example GPT-4. Although machine learning (ML) is often used now as a term to distinguish older types of AI (as in the AI so many were excited about previously), there is an overlap. GenAI focuses on content generation, ML on predictions; but you can use LLMs for some ML tasks such as classification. ChatGPT is also doing natural language processing (NLP), another older term that you will likely have seen bandied about. It may be hard to keep the terms straight, but who cares - it's shiny!

³ GPT-4 is the large language model available with the paid version of [ChatGPT](#). If you've only tried the free version, you've probably only used GPT-3.5. It's worth paying \$20 a month for the more powerful version. GPT-4 is generally considered the state of the art at the moment, but it's not the only LLM out there.

I'd tried GPT-3.5 when it came out in late 2022. It was fun, but when I ran it through a range of contracting scenarios, it flunked hard. GPT-4 was different. As just one example: I used to give a talk to students about AI where I would put up a clause, show how it could be translated into Welsh, have its entities identified, most usefully be classified by type, but my punchline was that it couldn't be interpreted by AI. I ran the clause through GPT-3.5 and it mangled it. But GPT-4? GPT-4 did a pretty good job with the interpretation, giving me a nice table of scenarios that it had made up and correctly showing the outcomes⁴.

I had two potential problems at this point⁵. The first was a bet I had already made with Casey Flaherty on Twitter, where I suggested that GenAI wouldn't displace even 5% of in-house and law firm time spent drafting, negotiating, or interpreting contracts within the next five years⁶. The second was that I had placed a far larger bet building Radiant Law around a people-first strategy⁷.

Neither of these bets is limited to whether GPT-4 is good enough. Better models will arrive within months, let alone years⁸. And clever people are doing clever unexpected things with what we have now⁹. This is consequently a hard space to predict.

But I'm going to try in this paper to make sense of how GenAI might affect contracting starting, where we should always start, with first principles and purpose.

⁴ Here are the two sessions: [GPT-3.5](#) and [GPT-4](#). GPT-4 made two mistakes in the first answer but saved it with the table. There are bonus law questions at the end (both versions mangle it, see below on not trusting it with facts).

⁵ Actually there was a third problem as I had earlier promised to shut up when AI interprets a clause, but hopefully no one remembers [that tweet](#).

⁶ The [bet](#) was for one drink (anywhere on the planet, flights included). Casey sportingly offered me the chance to back out when he later revealed that he had already had an insider pre-release glimpse of GPT-4 at the time of the bet. I declined, not only because I like having drinks with Casey and don't mind losing bets, but also because I may just win anyway, as I explain in this paper.

⁷ I like to call it a bionic lawyer, rather than robot lawyer, strategy. If we don't need lawyers and can just use AI (as suggested by so many breathless commentators), then I may have made a significant error. You may want to judge my arguments with an eye on the skin I have in this game.

⁸ Not just [GPT-5](#) (6, 7...), but OpenAI has just released a bunch of improvements to GPT-4 at the time of writing; Google is apparently going to release something called [Gemini](#); and the [open source LLMs](#) are doing surprisingly well. Chips are emerging that are [optimised for training models](#); [open source datasets](#) are appearing. Let's just assume that lots of things are probably going to be happening for the foreseeable and that this will remain a topic on the conference circuit for somewhat longer than blockchain.

⁹ I'm watching out for combining LLMs and symbolic reasoning, but in the meantime GPT-powered [autonomous agents](#) are interesting. The [scientists who hooked-up GPT to standard lab equipment and found a way to produce chemical weapons](#) also gave us an "interesting" moment. The sprinkling of AI over every existing legal tech tool has been, at least for me, less interesting.

Contracting needs to be better

One of the issues with the terms of the bet with Casey is that it put the emphasis on who or what will be carrying out, in future, the current everyday tasks of lawyers working on contracts¹⁰. More important is whether we are currently doing those tasks well.

As I argued extensively in my book¹¹, commercial contracting not only matters (life blood of companies etc etc) but we generally do it terribly. We send out unreasonable and near-incomprehensible terms, spend our time arguing about the wrong issues, and then fail to manage the outcome. Most importantly, contracts are not valuable per se; value is created through the commercial relationship, and contracting practices undermine the very commercial relationships we are creating.

The terms of the bet focused on three stages: drafting, negotiating, and interpreting contracts (not by accident, they are the obvious areas where GenAI might help). Let's dig further into the current problems in those stages:

Drafting

Contracts should be short, clear, reasonable, and relevant. We have extensive experience at Radiant of taking forty page contracts down to seven pages or fewer, without losing the substance. I described in the book an experience Radiant had where including just one (unnecessary and, worse, unreasonable) clause in a template caused contract negotiations to take over four times longer to close.

It's not only the terms themselves, it's how drafts are created. Although document automation is a well established solution for consistently and correctly drafting contracts in minutes, most legal teams are still not using it¹².

¹⁰ Another issue is that it's unclear how we will determine who has won. All my fault, I set the terms of the bet.

¹¹ [Sign Here: the enterprise guide to closing contracts quickly.](#)

¹² Document automation was invented in the '70s and commercialised in the '90s. There are now hundreds of products out there, as the next generation of legaltech vendors obstinately reinvent the past. Uptake is still low despite better interfaces. I see this as a wonderful (and tragic) example of the willingness of people to keep doing painful work, rather than take the time to fix a problem. We aim at Radiant to produce every contract using document automation. It's not only far faster, but removes many types of errors. Legal departments have all the incentives to use document automation, but seem to lack the budget, expertise (which is another way of saying lack of budget), or willingness to change. Law firms have an incentives problem: document automation reduces billable hours. Some law firms have still embraced it in the pursuit of higher value work, but they remain the exception.

Negotiating

WCC's regular survey¹³ continues to show that contract negotiations focus on what will happen if something goes wrong (liability, indemnities, termination etc) rather than ensuring that it goes right. The survey is an embarrassment to the entire industry. What matters¹⁴ is whether the customer is buying what it needs to solve its problems, and the supplier is selling what it can deliver (and the price is understood and right for both parties). But these key topics tend to get drowned out by the focus on worst-case scenarios, ironically increasing the chance of those scenarios coming true.

Meanwhile, the negotiations themselves are often counterproductive in style, with too many lawyers seeming to think that they are there to win (the wrong points), not start a commercial relationship off on a positive and thought-through footing.

Interpreting

If contracts were short and clear, they would be easier to interpret, but a pile of even simple contracts is still a large job for human reviewers. The problem is often framed in terms of finding the contracts, which is a good start, and extracting data points, but the issue is much worse: companies don't know how to make many of their contracts actionable (though they tend to be better with their standard sales contracts).

What is usually thought of as a need to extract data (which involves some level of interpretation) might more usefully instead be treated as a need to turn contracts into action points (also requiring some level of interpretation) that are then actually actioned. It is thus a people and process problem, which can't be solved with just a contract review project.

Still, many companies haven't even got to the extracting stage, and those that have often find they have to keep going back around when a new question arises. Meanwhile, contracts keep leaking value¹⁵.

The point

So, these key aspects of contracting are all generally done badly, and replacing humans doing them badly with machines doing them badly is not the "progress" that I want to be associated with.

¹³ World Commerce and Contracting 2022 [survey results here](#), but basically it's been more or less the same every time they've run the exercise since 2007.

¹⁴ As Jeff Carr, a voice of reason in this mad world, [keeps reminding us to little avail](#).

¹⁵ WCC famously identified a [9% bottom-line impact](#) from value leakage in commercial contracts. Your mileage may vary but let's stipulate that businesses often don't get as much from their contracts (or rather commercial relationships) as they might hope.

Contracting and contracts are a mess because of the focus on the wrong points:

- The drafts focus on the wrong points.
- The negotiations focus on the wrong points.
- The right points aren't actioned over the term.

I suggest, therefore, that the key question for whether (and how) to use GenAI in commercial contracting is whether GenAI can help with these issues.

To answer this question, let's start with some things you should know about GenAI.

Things you should know about GenAI

The following points are from my experience working with GPT-4, discussions with others, and much reading. I'm assuming that all current GenAI systems will produce similar, or worse, results but that may be incorrect. Models will also improve, but many of these issues are inherent in the technology. If someone is selling you something they claim is not hampered by the following limitations, you may want to dig deeper.

Generative AI generates text

As is blindingly obvious to anyone who hasn't worked with contracts: contracts are long texts, GenAI can generate long texts that read like they were written by a lawyer, so lawyers are done for.

I've already noted that contracts could do with being shorter, not longer, so there may be issues with using a system that produces text in bulk. I also noted that contracts need to be relevant (as well as reasonable), so the content of the text may be a little more important than our outside observer acknowledges.

GenAI works on a prediction model that generates answers that look right. That can include contracts. It's not a miracle; it's digested an awful lot of awful contracts in training, and is spewing out text that looks like the contracts it trained on. Moving past the problem with being trained on the dire drafting that is prevalent in our industry, the answer just has to look like what a human would produce. Whether the text works as a whole, is internally consistent, covers the objectives of the party, or randomly introduces issues that lead to lengthy negotiations is beside the point to the system. It produces language that looks like a contract, good luck with what it actually says.

Text that looks like a contract, and text that consists of only the right terms for your particular situation, are not necessarily the same thing.

GenAI sounds plausible

The models have been trained to give you plausible answers, and they are very convincing. The grammar is excellent, the tone feels right¹⁶, and each sentence generally parses. If you ask for a contract, the answer will look like a contract. If you ask for a limerick you will get a limerick¹⁷.

These plausible answers can lull you to sleep (more about this next) but they also trigger a problematic part of human reasoning. We tend to think by analogy. When we see a system doing something that we've only experienced humans doing, we may assume that it can therefore do other things that humans can do. This is dangerous, because it can and will fail at other tasks that "should" be trivial.

Humans are bad at checking walls of text

Some have suggested creating a first draft using GenAI and then checking the results. I think this is generally a bad idea, because humans tend to be unreliable at actually checking text that looks vaguely right. Contracts are highly technical documents, where the parts need to work with each other and things that are missing can be just as important as things that are included. We tend, in contracting practice, to assume that what we are reviewing is at least rigorous, even if it includes positions that we may not agree with.

Getting contracts right is incredibly hard, which is why templates are so important. You can't assume that humans will apply the same rigour to every document that crosses their desk.

GenAI doesn't reason (kind of)

When I first started thinking about AI, I was probably overly reductionist on the topic of reasoning. GenAI doesn't reason, in the sense that it can hold abstract ideas, understand how they fit together, and consistently draw the right conclusions about novel situations. But it does appear to reason, because it was trained on text imbued with reasoning. So it can look like it is doing exactly the same thinking that humans can do, right up to it failing in surprising ways.

The need for reasoning in contracting is a mixed bag. My point about the need to focus on the relevant points suggests that reasoning is a priority. But we humans often work on pattern recognition, and contracting is often capable of being reduced to patterns: are we dealing with situation X or situation Y? So GenAI may get to the right answer if you can get it to spot patterns, sometimes.

¹⁶ And you can change the tone through the prompt, a fun experience (the first few times).

¹⁷ Not necessarily a [good limerick](#).

It also can appear to reason about clause interpretation, sometimes correctly, and I described an example earlier. It is, however, inconsistent and gets worse if interrelated concepts are handled across the contract, rather than grouped in one clause¹⁸.

GenAI is random (kind of)

I keep saying “sometimes” because the system produces inconsistent results. If you’re in the story-telling business, consistency is a weakness. It’s boring. If you’re in the contracting business, consistency is a virtue. If you ask the system to produce a contract with the same prompt it will tend to give you a different one each time (even the inconsistency is inconsistent). You can reduce the “temperature”¹⁹ but it doesn’t fully solve the problem. By different I mean that any given clause is likely to vary enough to represent different positions on the same issue (if the issue is even addressed, in fact if the clause even appears again). Drafting with GenAI is a recipe for not only inconsistent contracts across your estate, but also a perpetual fountain of new fun topics to be negotiated with the other side. If you ask it for negotiation points, it will produce different results each time (usually).²⁰

GenAI is also a black-box system, meaning that you can’t get visibility of why it gave a particular response. You can force it to show its reasoning when answering a question, but the reasoning needs to be checked as well as the answer, and the reasoning is also meant to look good rather than be right - it may not stand up to scrutiny or actually support the answer²¹.

Another related issue is that the models are being constantly adjusted behind the scenes, without announcement. If an intermediary supplier is assuming that their fine-tuned prompt will always work, they may be surprised.

¹⁸ This is not just a function of how many tokens can be inputted (the context window). Some interpretations require seeing how concepts interrelate that are traditionally handled in very different places in the contract, may use unexpected terminology, and may interact to produce surprising results. This is harder to do consistently than when the topics are nicely grouped together. I expect this to get better.

¹⁹ Temperature is a setting where the lower the value, the more boringly consistent the results are.

²⁰ As an experiment, I asked GPT-4 in five different sessions to draft a two-way NDA using the same prompt (that asked it to try to be consistent across sessions). You can see the results here (at least two of them were one-way NDAs, it was hard to always tell what it was trying to do). Are you sure that you want it drafting for you? [Result 1](#), [Result 2](#), [Result 3](#), [Result 4](#), [Result 5](#).

²¹ And you can’t get the reasoning for the reasoning. There is a fundamental bootstrapping problem here.

GenAI can never be trusted with facts (on its own)

Because GenAI is producing answers that look right, with little regard for whether they are right, a GenAI system on its own can never be trusted with facts. You will have heard enough about, or experienced, the “hallucinations” problem to hopefully know that it is utterly unreliable.

That doesn't mean that a system incorporating GenAI cannot reliably produce the correct factual answer. People are figuring out how to connect GenAI to trusted sources of data and force it to use that data²². They are also building systems that can show links to the data sources so you can check them.

That's fine for getting the population of Paris²³, but it still leaves us with a conundrum where we don't have a golden source of truth. The good news is that contracts are rarely about either general facts or even specific questions of law (contracting is often light on law). Questions are more often along the lines of: what positions/rules apply to an issue under a particular contract? Is the issue even relevant to the contract, given the objectives of the party?²⁴

Prompting skills are both over and under-rated

Prompting, as in the questions you ask the system, is the perfect scenario for snake-oil sales techniques. By that I mean, the construction of prompts can be easily positioned as a mysterious topic, dependent on arcane secrets only known by the seller; and it is well-nigh impossible to prove such claims are wrong²⁵.

I recommend being highly sceptical of anyone telling you that they are in the magic prompt business. There is a lot of information out there on how to write good prompts, you can use GenAI to improve the prompts you ask GenAI, and no one yet has a proven path to the optimal prompt for any given problem (and better prompting will not resolve all the issues above). And no, prompt engineering probably won't be a career option for your children.

²² And OpenAI has just released a way that you can build and share custom versions (GPTs) of their model, with pre-provided inputs, such as data sources.

²³ 12,271,794 give or take.

²⁴ The first question requires actually checking the contract. The second is abstract enough that the knowledge might conceivably be held in a separate data structure. Doing this is non-trivial, but something we are working on at Radiant.

²⁵ A (hopefully) highpoint of nonsense was reached when [the sale of Casetext for \\$650m was justified in part by the eight magical prompts they had developed](#).

Having said that, the quality of your prompt will have a material difference on the quality of the output. It's worth reading up on what makes a good prompt²⁶. Starting with "You are a contracting expert" does seem to work!

Assume GenAI is trying to kill you

Admittedly a little overdramatic, and I'm not referring to a Terminator²⁷ situation. But the result of all of the points above is that if you don't assume that the system is trying to kill you, it will do dangerous things just at the point that you trust it. Please don't ever trust it.²⁸

Things you should know about humans

In case you take me as being fanatically "team human" over "team tech", I should note that the problems that plague contracting are not only the fault of humans, but I, and many of my fellow humans, bring all sorts of human flaws to how we work with contracts:

- **We are not Spock²⁹**: however selfless we may claim to be, our actions are often driven by our personal, rather than our organisation's, needs and we are rife with cognitive biases. We want to be heard, we want to be recognised, and we want to avoid being blamed. We over-emphasise risk, we are prone to complaining that others aren't accommodating our needs while ignoring theirs³⁰, and we are susceptible to anchoring³¹.

²⁶ [This Microsoft resource](#) is pretty good and they have [further guidance](#) on advanced techniques.

²⁷ The Arnold Schwarzenegger film franchise where an AI launches a world war against humans, and the source of such generation-defining quotes as "hasta la vista, baby" and "I'll be back".

²⁸ I asked GPT-4 to critique the "Things you should know about GenAI" part of this paper. Its conclusion was: "In summary, your piece paints a cautious picture of the use of GPT-4 for contract drafting, highlighting several inherent limitations and potential pitfalls. It emphasises the critical role of human oversight and expertise in leveraging AI for legal tasks, which is a well-grounded position given the current state of the technology. It's a sober and well-reasoned exploration of the topic." GPT-4 tends to err on the side of politeness, especially if you are polite to it; it appears to use how you write prompts as clues for the tone of the response. Anyway, following [Pascal's Wager](#), I try to be polite to our potential overlords.

²⁹ As in the Star Trek character from the fictional race of [Vulcans](#) who "are noted for their attempt to live by logic and reason with as little interference from emotion as possible". Trying here to balance out the previous [Star Wars references](#).

³⁰ I discussed this gem in [a previous post "On being reasonable"](#).

³¹ [Wikipedia's list of cognitive biases](#) is depressingly long and negotiations are basically applied psychology, despite academic writers' continued efforts to explain negotiations using [game theory](#). Note, however, that even though [Homo Economicus](#) is probably not coming back, many of the listed

- **We make mistakes:** our contracts are littered with errors and inconsistencies, and I suspect the fact that contracts rarely end up in court not only saves us transactional lawyers from the consequences, but allows us to be overly optimistic about the quality of our output.
- **We seem to prefer working hard to thinking hard:** many of the things I extol in my book as helpful for fixing contracting require doing real thinking up front (such as drafting a shorter contract or automating it) and we humans seem to prefer working to midnight rather than taking the time and effort now to fix it once and for all³². This is the best explanation I have for why contracting is improving so slowly. I don't, however, want to come across as puritanical: I would like doing the right thing to be easier.
- **We struggle to make tacit knowledge explicit:** much of the knowledge we want to convey around contracting is tacit, held in the heads of experts, and experts are generally bad at making their tacit knowledge explicit. We therefore rely on apprenticeship for training, probably over-rate experience (as we can't distinguish expertise otherwise), and tolerate inconsistencies from experienced lawyers because we don't know what should have happened. Knowledge management is usually an afterthought, and where it exists, it is most often an exercise in collating examples rather than extracting "the point".
- **We respond to incentives:**³³ we, contracting professionals, are also incentivised to make contracting complicated, because it justifies fees and salaries. It's not just that it is hard to make knowledge explicit and contracting simple, we are generally not incentivised to do so.

All of these loveable flaws suggest that we humans could do with some help.³⁴ If, despite the concerns I described in the previous section, GenAI can help us make contracting, and our lives, better, then I'm all for it.

cognitive biases are struggling to hold up under the [replication crisis](#), and biases can be helpful (they are there for an evolutionary reason).

³² I manually created several Radiant contracts last week because I **still** haven't got around to having them automated. Shame on me.

³³ My favourite quote is from Charlie Munger: "Never, ever, think about anything else when you should be thinking about the power of incentives." [I wrote about that in the context of partnerships and AI](#), one of several of my previous takes on AI that I am nuancing with this piece. Another example is [my post on AI and Satisficing](#) that also, on reflection, may have overemphasised reasoning as being binary rather than being on a spectrum.

³⁴ I think we can stipulate that GenAI is not going to fix incentives anytime soon, although law firms are muttering about needing to get rid of the billable hour because of AI. Given the long history of such muttering, with no change to date, I don't think we should take them seriously until they actually do it.

Suggested principles

Given the nature of contracting, and the fun issues with GenAI and us humans, here are a few suggestions for how to think about where GenAI may help, beyond filling a deep need for a new shiny thing:

Use deterministic systems where possible

It matters that contracts are precise. It also matters that your portfolio of contracts are consistent. You don't want every contract to take a subtly different position on an issue when considering what you can do at the portfolio level (for example, can we sell our business, or what are our obligations under our sales contracts?). Given this, I suggest that deterministic systems (i.e. systems that always produce the same predictable results from the same inputs) are preferable over systems that roll dice. We already have a number of deterministic systems that work well, such as document automation or workflow for well-understood processes.

However, these systems don't cover everything involved in the practice of contracting. Not all situations are predictable in advance or come up regularly enough to justify the investment. Some things can't be collapsed into if-then statements. And we should note that deterministic systems are hard to set up and maintain.

So, I'm not arguing that there is no space for GenAI, just that if a deterministic system will solve the problem, err towards using determinism rather than GenAI.

Harness iterative improvements

Back when I was a "real" lawyer doing deals, I observed that a novel clause I wrote usually needed to be improved over about three deals before I was happy for it to become a standard. It's not that it didn't work the first time, I just tended to find room for improvement over the next two, before the law of diminishing returns kicked in.

I've observed this pattern elsewhere in my work, and it has led me to value having a limited number of generic assets that keep being improved, rather than a large repository of examples. Not only do I get the benefit of improving the asset, rather than continually creating new things, but it also becomes easier to share with others (including because it doesn't require contextual knowledge of the situation where it was first used).

This observation, along with the challenge of checking GenAI outputs that I describe above, suggest that we will get more value from GenAI helping create improvable assets, than GenAI spewing out new deal documents.

Data structures matter

Given the value in improving things, rather than creating everything afresh, it matters where you keep things, including nuggets of knowledge. If you can't find something fast when you need it, then it has no value. If it doesn't live in a place where its relationship to other assets and nuggets is also represented, then it is harder to use.

This suggests that knowledge management is becoming even more important.

I have [written elsewhere](#) about how documents are coffins where information goes to die, and document management systems are therefore graveyards. GenAI may help us find example contracts faster, but it will remain hard to infer the undocumented background to the deal that led to the outcome. I suspect that we will be better off extracting the key points when we do the deal³⁵, and creating places for them to live. Yes this requires discipline, but we've found ways to do this at Radiant, so it's not impossible.

As luck would have it, for the year before GPT-4 dropped, I happened to have been working on these kinds of data structures³⁶, and my conclusion is that they are not only invaluable in conveying knowledge, but that without them it's going to be hard to get good results out of GenAI.

Prompt your brain not the machine

I've found it most useful to treat GenAI as a tool to prompt my brain. If I ask it to give me a list of points, it seems to trigger my thinking and I can quickly add, remove and clean the list up. It's not only far faster than starting with a blank piece of paper, it's so much easier (thinking IS hard) and leads to a much better result.

What I've found, therefore, is that with the right data structure (where I can save the output), GPT is a fabulous tool to stimulate my brain and help solve the gnarly problem of making tacit knowledge explicit. Humans may be bad at dealing with walls of text, but asking for bullets as a starting point for thinking works brilliantly for me.

It may also have a wider impact: we have more than our fair share of smart people in our industry, but smart people tend to be wary of straying outside the areas where

³⁵ This is not just another clause database, although that may be one element. We've found that as we build richer data structures, it becomes easier to collect richer knowledge. The main problem that lawyers seem to face is not knowing where to put their nuggets, which seems to be even harder than articulating the point. Of course, incentives matter here, too. The emphasis we put on collaboration and sharing at Radiant helps, and is a stark contrast to the average law firm.

³⁶ I should write about this, but for now I would note that no/low code tools like Airtable have had a far larger impact at Radiant in making sense of the world than GenAI has. What GenAI has helped with is populating the data structures, but I've been at pain to make sure that everything going in is checked.

they know what they are doing. But many problems in business life are not so complicated with some basic understanding of the domain³⁷ so it's worth exploring whether you really need all those external consultants.

Automate everything you can

I am aware that I strike some as a curmudgeon when it comes to legal technology, but I really would like us to automate everything that we can³⁸.

I want widespread automation, not because I believe that computers are now brilliant, but because I believe humans are, and have always been, brilliant. I want humans to spend their time playing to their strengths, rather than performing repetitive tasks that computers do better.

I said at the beginning that contracting is not about contracts, it's about relationships; and humans create relationships. I want the people working in contracting to be spending their time thoughtfully creating those relationships, doing the reasoning that GenAI still struggles with, and figuring out what the point is.

How might we use GenAI in contracting?

So we seem to have a system on our hands that does things that shouldn't be possible, but does them haphazardly and in a way that can easily lure us into a false sense of security. It has lots of issues, but so do humans.

GenAI also has incredible strengths:

- it can respond on a huge range of topics, apparently unhindered by the strict boundaries that come with deterministic systems,
- it can spark thinking if used as a prompt,
- It can vary the tone of text and its grammar is generally excellent,
- it can summarise pretty well, and is going to get better at interpreting,
- it can give a starter for 10 for many documents (although that isn't always helpful),

³⁷ There are limits here, playing in strategy (the brilliant [Roger Martin](#) aside) is probably safer than medical diagnosis. Beware though that the areas where this applies best are probably those that are, on final analysis, [more bike-sheds than nuclear power stations](#), so more opinions on simpler stuff may not add as much value as we might hope.

³⁸ With some caveats, of course, such as first figuring out whether the activity is adding value, getting it right manually first, paying attention to the cost of change which may be far higher when a process is automated etc. But still, all things being equal, I like automation.

- it can manipulate text and will likely get better at making changes to documents (even if it remains haphazard at figuring out what changes are necessary),
- it can give you the ability to experiment with AI without needing the technical expertise that was required to date (this is huge), and
- more generally, it can do so many things better than me where I am a novice (and I am a novice at most things in life)³⁹.

It also feels like this is only a start. As I mentioned earlier, not only will GenAI get better (although that is not the same as a pathway to general AI), but smart people are experimenting and will keep coming up with novel use cases. I don't think this is the holy grail for AI, I don't think it will or should replace all legal tech, but I am cautious about declaring the limits of where it might be used, despite the concerns I have raised.

With those thoughts in mind, here are some use cases where we have found GenAI helpful, or I suspect it will add future value⁴⁰.

Creating assets

I've had some success with throwing in examples of deals and asking for framework templates back (you need to apply a lot of clean-up and judgement on the way but it helps me see the wood for the trees), and to a lesser extent, figuring out which questions to ask for document automation.

Where I feel I've really struck gold is prompting my brain when creating knowledge resources for "things you should know about X". I need a place to put them, though, as explained in the data structure section above⁴¹.

³⁹ I have written a LOT of code using it. Code is a nice use case, because it can be easier to test code by trying to run it than, say, factual output. For my use cases, disposable code has been good enough. I would still want a developer rewriting it if it matters. It's a lot easier, by the way, to underestimate what other people do, as lawyers have come to experience in this debate, than to expose oneself to the complexity in their job. I don't think lawyers are going to be replaced, and I definitely don't think that developers are.

⁴⁰ The list would be a lot longer if I included all the AI tools that are popping up in the current bubble. There are even some declaring game over for pretty much all existing legal tech, with it being replaced by OpenAI. My conversations with a number of the real thinkers in legal tech suggest a shared opinion that valuable use cases may be limited for contracts. However, everyone is under pressure from customers and investors to sprinkle AI fairy-dust over their products, so it's going to be a confusing picture for some time. And I'm quite sure (a known unknown) that use cases will emerge that I haven't thought of.

⁴¹ Surprisingly, at least to me, was that we have had more impact in populating our data structures experimenting with how to show lawyers slices of the data in a way they can quickly fill in the gaps, or improve the wording, than using AI.

We've also had pretty good results doing boring heavy-lifting with mechanical parts of asset creation (e.g., classifications). There may be better specialised tools out there⁴², but the ability to get something going quickly and cheaply (often via the API) and its incredible flexibility, has been very helpful.

I suspect our biggest short-term use case will be accelerating building assets.

Drafting

You've likely gathered that I'm not a fan of drafting contracts using GenAI (although it can help with creating templates). We have found one use case, for parts of a contract that are normally so wildly divergent that they can't be solved fully with document automation, and have a prototype up and running.

Generally, though, please use document automation.

Improving negotiations

Ever since seeing GPT in action, I've worried about its potential for weaponising negotiations. I suspect that we are going to see more random points raised that add little to no value as GenAI is asked to review contracts.

I do see a few areas, though, where it might help:

- a thoughtful person, faced with a novel contracting situation, may find GPT useful to stimulate thinking about what might be the points that matter,
- more generally, I've suggested that negotiations are as much about psychology as substance and a bit of AI magic applied to the tone of comments, emails and changes may help make things more palatable,⁴³
- We've already got tools that connect clause classification (which ML has gotten pretty good at) with rules to spot issues in contracts (basically whether a clause you want is missing, or a clause you don't want is included). I imagine these tools, including the application of changes to the contracts, will improve. But this is only going to be helpful where the right points are raised. This requires either (a) someone actually thinking about the deal in front of them, or (b) a thoughtful playbook being created in advance, customised for the deal type and your organisational needs, that controls the changes being applied⁴⁴.

⁴² Noah Waisberg, who knows a bit about this area, wrote an [excellent analysis](#).

⁴³ Grammarly was pretty good at this already.

⁴⁴ Another example of why we need to do the work upfront if we want to take advantage of these systems.

Negotiations are so often random and awful now that I'm not sure whether GenAI is going to help or make things worse. I fear worse.

From our perspective at Radiant, we will experiment but it's going to be pretty marginal because we are already operating at 90% half-day turnarounds on contracts (either drafting or responding to red-lines) without much, if any, highfalutin AI. Any improvements are therefore going to be relatively marginal and, more importantly, beside the point: the game is to send out such deeply reasonable (and short, clear, and relevant) terms that we get rid altogether of negotiating the parts of contracts that lawyers traditionally care about (negotiations of the "commercials" can be very valuable).

Which takes us to another area: we've built our own tools for figuring out where contracts are being over-negotiated, but such tools are not widely used. If GenAI's analytical capabilities can help more teams figure out where they are shooting themselves in the foot with their terms, that would be a major win.

Reviewing signed contracts

We have only experimented in this area, rather than trying to use it at scale, but I suspect that contract review projects are going to succumb even further to technology (although finding the contracts may remain hard for some companies!)

I am even more convinced (I suggested this in the book), that the best strategy generally is to focus on (a) getting contracts in one accessible digital space, and (b) improving your processes to manage your contracts and make them actionable (including handovers between the negotiating team and operations). However tempting it is to put large contract management systems in place or to review every contract for all conceivable questions:

- this is the time to hedge bets on what will be the best future technology and large complex systems (CLMs, I'm looking at you) usually end up getting in the way rather than solving everything,
- the question you will need to answer in future will almost inevitably be different to what you expect (and the tech is coming to make it easier and easier to answer future questions),
- you often don't need perfect answers to questions anyway (odds are that any particular contract based on your template won't have a change to a particular clause that you are interested in, so you can reason probabilistically with sampling), and
- the hardest part is to get your organisation to actually do things - this is a human/process problem, so a focus on making your contracts actionable and

making sure that people know what they need to do is more important than pursuing a digital nirvana.

Everything else

One of the problems with buying technology is that it is easy to focus on the moments when it might have been helpful in the past, while overestimating how often those moments occur.

Most contract lawyers don't spend most of their time drafting, negotiating, or interpreting contracts. There's a remarkable amount of "stuff" in all of our days. And even when working on contracts, if you've done the modicum of hygiene I suggest in my book for standardising, automating, and improving what you can, most of what is actually going is weird edge cases, aligning colleagues, and wrangling organisational nonsense.

With that in mind, figuring out how to work well with GenAI outside the strict parameters of what is "contracting" may have a far greater impact on what you can achieve in a day than applying it to contracting itself. Although I have given only limited use cases where I think GenAI will help with contracting, it may still add lots for you as a tool. The only way to find out is experimenting. No one is an expert in how you should use GenAI: these are not only very early days, but you are always going to be the best person to figure out how it may help you⁴⁵.

Why I may still win the bet

Let's return to my bet with Casey: that GenAI won't lead to more than a 5% reduction within five years in the time spent drafting, negotiating, or interpreting contracts.

I really don't know what the outcome will be, but the following has to be true for Casey to win:

- **Useful:** GenAI is not only relevant to drafting, negotiating, and interpreting contracts but also materially reduces the time taken to do the work.

⁴⁵ Centaur Chess, which was invented by Gary Kasparov after he became the first chess world champion to lose to a computer, may be relevant here. The game is played by humans plus computers against other humans plus computers. Here's what Kasparov learned: *"The surprise came at the conclusion of the event. The winner was revealed to be not a grandmaster with a state-of-the-art PC but a pair of amateur American chess players using three computers at the same time. Their skill at manipulating and "coaching" their computers to look very deeply into positions effectively counteracted the superior chess understanding of their grandmaster opponents and the greater computational power of other participants. Weak human + machine + better process was superior to a strong computer alone and, more remarkably, superior to a strong human + machine + inferior process."* The Indian grandmaster Viswanathan Anand said this about the experience: *"I think in general people tend to overestimate the importance of the computer in the competitions. You can do a lot of things with the computer but you still have to play good chess."*

- **Adopted:** even if it is useful, it still has to be widely used by in-house teams and law firms.
- **Demand stays constant:** putting aside there may be more deals in the future (likely as demand seems to be generally rising but that wouldn't be fair to count), could making things easier increase the demand for it?

I think I have a fighting chance on all three fronts.

Usefulness

As I've discussed in this paper, I think GenAI will have marginal if any usefulness for drafting, may increase the amount of time spent on negotiations if mis-used, but will make it far faster to make sense of your signed contracts. However, I'm not sure that much time is currently spent reviewing and interpreting deals (other than relatively rare review projects), because reviews are so painful right now that everyone is used to just ignoring what was written down. So there isn't much time to reduce in the area where GenAI will probably have the biggest impact.

Adoption

In-house and private practice lawyers have been pretty bad at adopting technology that helps. This is partly explained by the incentives that apply to private practice⁴⁶, but I'm still astounded by how little document automation (which actually works) is used by in-house teams. Even though AI is going to find its way into all the tools we use day-to-day, most features already in Word, Outlook etc are ignored. I'm not convinced that five years is long enough to make an impact, even where GenAI offers a clear advantage.

Demand

Ironically, Jevons Paradox⁴⁷ suggests that even if GenAI is useful and adopted, reducing the time/cost of doing the activity may increase the demand so even more time is spent by the team doing the activity. I suspect we are going to have a lot more questions in the future about what's in our contracts. Whether that translates into better contract management is moot.

⁴⁶ Cough, billable hour.

⁴⁷ From ChatGPT: "*Jevons Paradox states that as technological improvements increase the efficiency with which a resource is used, the rate of consumption of that resource tends to increase rather than decrease. This is due to the fact that the increased efficiency effectively lowers the cost of using the resource, which can lead to an increase in demand that outpaces the gains in efficiency. In simpler terms, when something becomes cheaper to use (often due to technological advancements), people tend to use it even more, rather than less.*"

Conclusion

GPT-4 is amazing, but the closest we have to a silver bullet in contracting remains making contract templates short, clear, reasonable, and relevant. AI might help you do this, but it still requires human judgement calls and the very human ability to navigate corporate decision-making - it's always about relationships.

I'm looking forward to my drink with Casey, whoever is buying, but in the meantime here is ChatGPT's argument for why I might be wrong:

"Despite the scepticism surrounding the impact of GenAI on legal practices, there is a compelling case to consider that Casey might be right in his bet, suggesting that GenAI could significantly reduce the time spent on drafting, negotiating, or interpreting contracts within the next five years. The rapid advancement and increasing sophistication of AI technologies like GPT-4 point to a future where AI could handle a broader range of complex tasks with greater accuracy and efficiency. As AI systems continue to learn from vast amounts of legal data, they might become capable of performing intricate legal analyses, generating more precise contract drafts, and even predicting negotiation outcomes, thereby streamlining the contracting process significantly. Furthermore, as the legal industry becomes more familiar with and trusting of AI capabilities, adoption rates could surge, leading to a widespread transformation in how legal work is conducted. This scenario would not only validate Casey's perspective but also challenge the current belief that AI's role in legal practices is primarily supportive rather than transformative."⁴⁸

⁴⁸ I then browbeat it into choosing a side: "Based on the current trajectory of AI development and considering the conservative nature of the legal industry's adoption of new technologies, the preponderance of evidence suggests a more cautious approach. Therefore, the choice would be "Alex." This reflects the view that while AI will significantly impact the legal field, it might not lead to a dramatic reduction in time spent on core legal tasks such as drafting, negotiating, or interpreting contracts within a very short timeframe." Good computer!