NOTES ON THE STYLE OF THE LAW
Using AI to determine 'plain meaning' is foolish

by

ELIJAH Z GRANET

P LAIN meaning is one of those elusive abstracts that occupies the irreality of the Legal Village, where the Reasonable Man on the Clapham Omnibus passes it daily on his commute to work as a reasonably circumspect and competent doctor. It is an elusive and tricky concept, whose name belies the challenge of finding the artificial balance in the judicial consideration of words and phrases.

Given this, some are tempted to try to outsource, at least in part, the hard work of the deduction of plain and ordinary meaning to technology, particularly the (perhaps inaptly named) new developments in artificial intelligence (AI). One recent call for such a rethinking came in the USA, where, in a case hinging on the meaning of the word 'landscaping', Newsom J, of the Eleventh Circuit,[1] filed a lengthy concurrence arguing that AI could be of some use.[2]

In short, the case involved, albeit tangentially, on whether allegedly tortious conduct (the installation of a trampoline) was covered by a general insurance policy covering the appellant's work in 'landscaping'.[3] The case did not turn on this point,[4] but Newsom J, in what His Honour admitted was an eccentric move,[5] nevertheless went on to make a lengthy case for at least the consideration of the use of AI perhaps inform the deduction of interpretive meaning. Note the hedging of the argument; His Honour acknowledged that even this 'modest proposal' was likely to be condemned as heresy and was careful

---

1 The United States Court of Appeals for the Eleventh Circuit hears appeals from federal district courts in the States of Alabama, Florida, and Georgia.
2 *Snell v United Speciality Insurance Co* № No. 22-12581 (11th Cir, 28 May 2024)
3 *ibid* (Opinion of the Court) at 2
4 *ibid* (Opinion of the Court) at 15, *per* Branch J
5 *ibid* (concurrence) at 1, n 1, *per* Newsom J

to limit the scope of the suggestions.[6]. This intellectual humility is to be praised,[7] but, for the reasons I will set out, His Honour is nonetheless gravely mistaken even in these circumscribed proposals.

Newsom J set out the traditional manner in which a judge might canvas sources for starting points on ordinary meaning. First, there are the various dictionaries of the English language.[8] This proved, as consulting dictionaries so often does in quests for plain meaning, less than satisfactory. Newsom J then found a much better way of arriving at a conclusion of the plain and ordinary meaning: looking at the evidence of the case (photographs of the trampoline installation) and using his judicial common sense and experience to arrive at the conclusion that this did not fall in the ordinary meaning of landscaping.[9]

This really ought to have been the end of the matter. A judge brings to the bench her experience and common sense and the utilisation of those invaluable tools is an accepted and institutionalised part of the legal process.[10] His Honour's observation of the photographic evidence on the record and the conclusion made, based on common sense and experience, that this was not within the plain and ordinary meaning of 'landscaping', was an entirely coherent judicial decision.

Yet, Newsom J lacked confidence in asserting the traditional rôle that common sense and experience play in determining ordinary meaning. His Honour wanted detailed articulation as to '*why*' (emphasis in original) this was the case.[11] Here, the problems begin. The legal questions of what is reasonable, what is the plain meaning, what is a preponderance of the evidence, where reasonable doubt lies as a threshold, are, no pun intended, matters of judgment. They require a mature and experienced jurist to use experience both inside and outside the law to come to a rational conclusion, even if it is sometimes hard to elaborate on why the conclusion is so other than *res ipsa loquitur*. Unfortunately, this reality of judicial decision-making gave Newsom J the 'willies'.[12]

Consequently, working alongside His Honour's law clerks, Newsom J decided to consult

---

6   *ibid* (concurrence) at 1

7   Indeed, the courts of the common law world would be much better of generally if other judges showed similar awareness of the limitations of their innovative proposals.

8   *ibid* (concurrence) at 5f

9   *ibid* (concurrence) at 6–8

10 See, *eg*, *Ashcroft v Iqbal* (2009) 556 US 662, 679, *per* Kennedy J ('Determining whether a complaint states a plausible claim for relief will […] be a context-specific task that requires the reviewing court to draw on its judicial experience and common sense.')

11 *ibid* (concurrence) at 8

12 *ibid*. Newsom J has a colloquial style which is rather different to the style of this publication, but His Honour has cultivated it and it works as an effective and persuasive judicial voice (even if here persuasion failed). The use of the more demotic phrasing is thus not an error but rather the deliberate and reasoned choice of an experienced legal writer. Once again, what matters most in style is not the decisions made, but rather that the writer has good and cogent reasons behind her decisions and works to accomplish those stylistic choices with skill and practice.

---

various popular online large language models of AI.[13] The 'ChatGPT' model put out an anodyne, though not hallucinatory, answer. This answer, though, did not go much beyond the few sentences of reasons Newsom J himself could have easily come up with to articulate why the photographs did not depict 'landscaping'.[14] Indeed, given Newsom J's skill as a writer, it was rather a sight below what His Honour could have produced with little effort.

Again showing praiseworthy humility,[15] Newsom J recognised that this risked simply searching for confirmation of His Honour's conclusions. Instead, His Honour chose to articulate a tepid, but open case for the potential further use of large language models as one of many tools in determining plain meaning. Yet, even this moderate position, which is really nothing more than the endorsement of further investigation is mistaken in its premises.

In favour of the potential benefits of the use of large language models, Newsom J first noted that they are trained on ridiculously large corpuses of ordinary language, including the intensely plain and ordinary speech of the Internet.[16] This is a weak point, because human beings, too, are trained in the ordinary use of their language (and especially in the spoken language, which can differ rather radically than anything in the written corpus of a large language model). The notion that a large language model is in better touch with the use of ordinary words than a judge or lawyers both underestimates jurists and overstates computers. The added value here is nonexistent, especially because unlike the black box of large language models, a judge using common sense and experience can point to the reasoning, where it exists, behind a conclusion and equally where it is simply a trite matter of common sense. A large language model is not aware of why it reaches a conclusion, nor can it distinguish between a hallucinatory correlation and an actual reason based on data. Even in the hedged proposal of Newsom J to consider using large language models as merely one of a number of inputs, it makes no sense to admit into consideration the worst sort of hearsay, an unreliable and often unreplicatable output (with unknown levels of stochastic variation).

This leads to the problems with the second argument in favour. Newsom J considered the probabilistic models underlying large language models and the degree to which this statistical model can provide an objective or at least relatively reliable guess as to how ordinary people would use language.[17] Here, with the greatest of respect to His Honour, this point is rather undermined by the fact that the evidence for the broad and imprecise

13 Given that LLM is traditionally used for the degree of Master of Laws (a degree to which your correspondent has himself been admitted), this publication declines to use that abbreviation for large language models. Readers are welcome to suggest alternate abbreviations for these models by comments or correspondence.

14 *ibid* (concurrence) at 8

15 I stress this is not empty flattery; His Honour's scepticism and methodological rigour towards the use of AI ought to be a model to other judges. In particular, the present Master of the Rolls on the other side of the Pond could learn a thing or two from this example…

16 *ibid* (concurrence) at 11–14

17 *ibid* (concurrence) at 14f.

claims made about large language models is itself a law review article.[18] This is a matter that calls not for judicial takeaways for the popular science version that is accessible to lawyers, but rather for expert evidence by those who understand the mathematics and architecture of large language models.[19] Newsom J's firm conclusion that large language models are 'high-octane language-prediction machines capable of probabilistically mapping [...] how ordinary people use words and phrases in context' is both scientifically imprecise ('high octane' is a metaphor of little use) and without evidential foundation. This lays out one of the great pitfalls of large language models and similar new technologies—that the fancy bells and whistles impress people into granting them a credibility and praise which they have not earned.[20]

Newsom J's arguments as to costs don't really persuade; the price is irrelevant if the method is not reliable. His Honour's next point is that compared to dictionary definitions, which may be put in by a few harried lexicographers and should not be given over-weighty authority.[21] This is certainly true. However, Newsom J then attempts to claim that large language models are more transparent and good at accurately predicting normal language. These are nonsense claims with no evidence to support them; large language models are indeed rather opaque and their inputs do not necessarily reflect ordinary language (particularly when so much of ordinary language is not written). As a supplementary point, His Honour bemoans that choosing which dictionary definitions to favour involves discretion and that discretion is hard to explain.[22] That may be true (though I think it to be no bad thing), but the large language model is incapable of reliably and replicability explaining why it 'chose' a particular method. There is no advantage here and some considerable disadvantage.

Finally, Newsom J argues that large language models may have advantages over corpus linguistics or surveys in determining ordinary meaning.[23]. This may be true, but strikes me as irrelevant, given that both those other methods also have enormous drawbacks.

Newsom J also, again to His Honour's credit, went through some of the downsides to the use of large language models, and makes clear (again) that this is not a call for the elimination of human judgment but merely considering using large language models as one input into the judge's final decision. That is all well and good, but even in this limited capacity, Newsom J is wrong to thing that large language models 'have promise' in this particular area of legal work.[24]
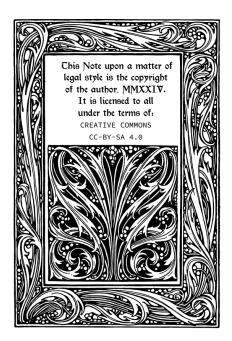
Ultimately, the lengthy and thoughtful discussion given by Newsom J is a long way of

18 Yonathan A Arbel & David A Hoffman, 'Generative Interpretation' (2024) 99 NYU L Rev (forthcoming)
19 The fact that even those who design such models do not fully understand the manner in which they work is indicative of why this is an unreliable path.
20 One such person is plainly, on this side of the Pond, the present Master of the Rolls…
21 Snell (n 1) (concurrence) at 18
22 *ibid* (concurrence) at 18f
23 *ibid* (concurrence) at 20f
24 *ibid* (concurrence) at 29. The technology certainly has promises in *other* areas, but this Note is concerned with the specific application to determining ordinary meaning.

expressing fear at relying on judicial common sense and of searching for some more objective input than the common sense and experience of a judge with the English language. This is all misplaced. Newsom J's initial conclusions from examining the photographic evidence and considering it in light of His Honour's own experience with the ordinary use of language was ample evidence. Combined with a few sentences exploring the meaning of 'landscaping', and indeed without any need to even look at dictionaries, this is the strongest way to establish ordinary meaning. The stochastic output of a large language model, which might produce two different answers at different times to the same prompt, can only serve to distract from the exercise.

This is particularly clear because the search for ordinary meaning, as I discussed in the first paragraph, is not actually a search for the objective way in which people use language. It is a legal construct of the juristic mind which, like many such legal constructs, intersects with reality, but does not perfectly correspond to it. The ordinary meaning of a word when judicially considered can be and often is rightfully quite distinct to the ordinary meaning of a word when demotically considered. In this sense, Newsom J's quest reflects the treason of language. The lawyer's phrase 'ordinary meaning' is a concept shaped by precedent into a divergent but closely parallel lineage to the layman's phrase 'ordinary meaning', just as the lawyer's use of 'reasonable' is in a divergent but closely parallel lineage to the layman's use of 'reasonable'. To try to seek some objective or scientific answer in this process is to turn against the way in which legal history has shaped the precise contours of that which we in the law misleadingly call 'ordinary meaning'. (Much the same can be said, and indeed has been said, of the use of dictionaries in this exercise.) Even if we could eliminate all the technical problems I have discussed involved in the use of large language models as inputs, going beyond the already shaky use of dictionary definitions as evidence of that which we call 'ordinary meaning' is a futile and misguided exercise.

For more Notes on the style of the law, visit legalstyle.co.uk

Published
by

GRANET
PRESS
LIMITED