# Supporting responsible AI

Submission from Interactive Engineering Pty Ltd – activesemantics.com

Generative AI using LLMs is inherently unreliable. The method uses word association, and has no understanding of what the text means.

How to regulate it?

If you insist the output must be accurate, then the technique can’t be used.

If you allow the technique to be used, you can’t regulate it, because there is no element inside the LLM which would have any understanding of the regulation.

Google on their version:

*You can think of an LLM as a prediction engine. When given a prompt, it generates a response by selecting, one word at a time, from words that are likely to come next. Picking the most probable choice every time wouldn’t lead to very creative responses, so there’s some flexibility factored in.*

When a word has many meanings – “set” has 72, “on” has 77, “run” has 82 – this comes across as an extremely stupid thing to do. But people will buy it for the novelty, so should we let them?

All sorts of effects will follow – group think, even more resistance to new ideas because they are new and not in the model, a community of blissful ignorance.

If you want to regulate something, it would help if the “thing” had an understanding of meaning. This is the area of AGI (Artificial General Intelligence), which does “understand” meanings, and how words are clumped. Examples:

A chess set, a movie set, he dropped a set at Wimbledon, the rain set in, he is set in his ways, it is set in stone.

An imaginary flat surface – the “flat surface” is imaginary. The adjectives are order-dependent, not like “a large black car”, where the adjectives are independent, even if presented in a conventional order. Words are clumped to form complex objects, and are further clumped to form even more complex objects.

The first problem with regulation is, how do we manage the collaboration among people from many different specialties. A good example was the failure of collaboration between economists and epidemiologists during Covid – almost no common vocabulary or intent. The more different specialties brought in, the worse it gets. One suggested list had lawyers and ethicists, engineers, scientists, and Uncle Tom Cobley and all. Humans have a Four Pieces Limit for their conscious mind, so it is not possible to share a concept that touches many different areas across many different specialties – the result is they would not agree on anything, or they al have different interpretations of what the regulations mean.

We have been working on this area of AGI – Robodebt was a good example of allowing lawyers, programmers, social workers, and those with hidden agendas loose on something. Another example is large, complex specifications, with many different specialties involved. The result tends to be something that nobody understands, and has the predictable result of an expensive schemozzle.

To put it politely, you are going to need help from AGI to create meaningful regulations that a machine can work to.