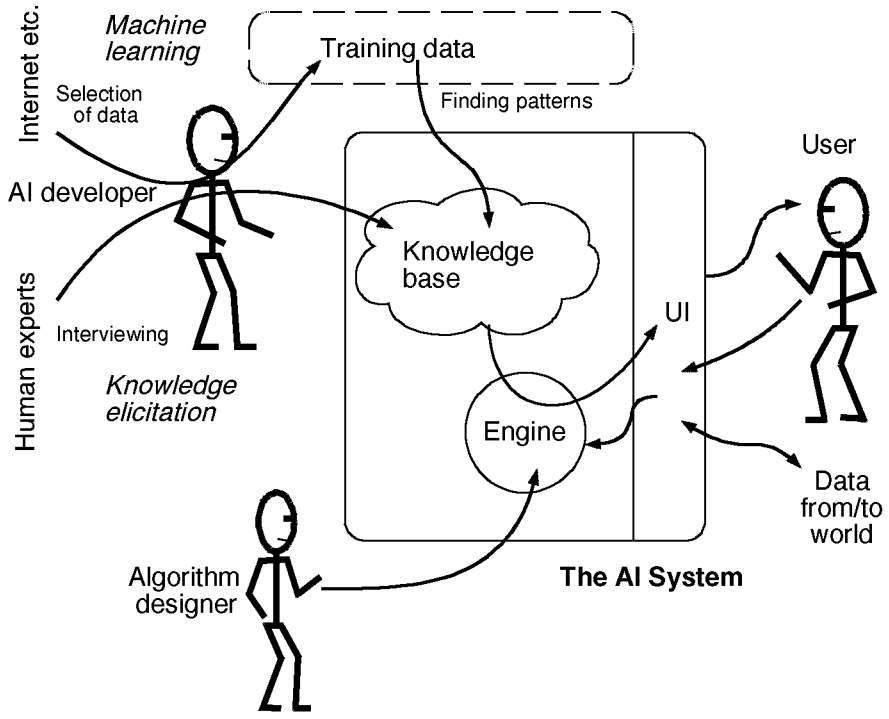
3 of 3. How Well Can AI Work, and Why?

Andrew Basden

My first blog [FOOTNOTE 1] showed a way to address the AI question "AI = Human?" more fruitfully than is usual, by reference to multiple aspects of reality. My second blog [FOOTNOTE 2] described how AI systems (like X-ray analysis or ChatGPT) work in general, and the inescapable role of humans in AI. AI systems comprise three main components, with different human roles for each, as depicted in the following Figure reproduced from Blog 2.



In the 1980s, I, as AI developer, encapsulated knowledge by interviewing experts and expressing their knowledge in a computer language. Today, with machine learning AI (MLAI), the developer selects vast amounts of data to train the knowledge base to recognise meaningful patterns. Read Blog 2 for more.

With this understanding, we can address the five remaining questions listed in those two blogs.

--- What Makes AI Capable?

The capability of an AI system comes mainly from its knowledge base encapsulating laws and information that are meaningful in aspect(s) of reality relevant to its application: spatial aspect for Chess AI, kinematic for automated cars and lingual aspect for ChatGPT, for example. [FOOTNOTE 3]

How can ChatGPT write essays, for example? ChatGPT analyses user's instructions or questions, and generates the text of the essay. Both operate according to the laws of the lingual aspect, which are encapsulated as host of probabilities of which phrases and words tend to follow which [FOOTNOTE 4]. This was constructed by ChatGPT reading vast amounts of Internet content, which, of course, emerged from humans functioning in the lingual aspect (consciously or subconsciously).

Table 1 gives Dooyeweerd's fifteen aspects, with their laws and some examples applications for each that are mentioned in these blogs.

	Aspect	Laws to do with ...	Applications mentioned
'Successes'	Quantitative	Arithmetic	-
	Spatial	Spatiality	Chess; X-ray analysis
	Kinematic	Movement	Automated cars
	Physical	Energy, forces, causes Physics, chemistry	Molecule design
	Organic / Biotic	Life functions, cells, organs, Organisms, Ecosystems	Face recognition
	Psychic / Sensitive	Stimulus-response	Voice extraction
'Human' aspects	Analytic	Concepts, analysis, logic	Business analysis
	Formative	Design, planning, forming	-
	Lingual	Signifying, speaking, writing	ChatGPT
	Social	Associating, institutions	-
	Economic	Resources, frugality	(Business analysis)
	Aesthetic	Harmony, enjoyment	-
	Juridical	Rightness Reward, punishment	-
	Ethical / Attitudinal	Self-giving love	-
	Pistic / Faith	Belief, commitment Ultimate meaning	-

In fact, most AI applications encapsulate laws of more than one aspect. Chess AI, for example, encapsulates some laws of human strategy (formative aspect). Assisting ChatGPT's main operations are a database of relationships among people (social aspect), and knowledge of images (spatial, psychical)[FOOTNOTE 5].

But why does AI make mistakes, such as in automated cars not recognising a cyclist pushing a bicycle. or ChatGPT offering its famous "hallucinations"?

--- Why Does AI Go Wrong? (Q4)

There are several reasons AI goes wrong. One is errors in user input or world data. Three others arise from deficiencies in the encapsulated knowledge.

1. Erroneous knowledge. Because human writings from the Internet contain errors, ChatGPT 'learned' erroneous patterns that generate "hallucinations". Also, since its word associations are probabilistic, it sometimes selects inappropriate ones.

2. Missing knowledge: minor biases. Tacit knowledge and rare exceptions are often absent from a knowledge base. In knowledge elicitation, a good analyst will deliberately seek these out but MLAI learns patterns statistically. There is often not enough training data to learn rare patterns reliably, such as cyclists pushing rather than riding bicycles.

3. Missing aspects: major biases. Omitting a whole aspect omits a whole swathe of knowledge that is meaningful in that aspect. Whole aspects might be missing if the AI developer fails to recognise their relevance and so does seek them or provide training data about them. This becomes problematic when AI is used in different contexts. Most training data for ChatGPT was written by

affluent people in the Global North, in which some aspects important elsewhere have been undervalued.

It is the AI developer who is responsible for ensuring high quality knowledge bases. This becomes more challenging in later-aspect applications.

--- In Which Applications Can AI Work Well? (Q5)

In which applications AI is likely to work well (now and in future), can be understood via aspects. The laws of earlier aspects are easier to encapsulate in a knowledge base reliably. for two main reasons. One is that the laws of earlier aspects are more determinative so that, for example, 3 + 4 is always 7 (law of quantitative aspect), whereas a question might be answered is several different ways (lingual aspect).

The other is that the laws of earlier aspects act as a foundation for those of later aspects, so, in principle, encapsulating knowledge of later aspects requires us to encapsulate laws of all earlier aspects too. Laws of physics depend on three earlier aspects, those of lingual, on eight. Moreover, the middle aspects of human individual functioning are influenced by later aspects too, which can also need encapsulating (e.g. ChatGPT's social database).

Therefore AI tends to work more reliably, and have more successes, in applications governed by the earlier aspects, than those governed by later aspects (see Table 1). X-Ray analysis (spatial aspect) is more reliable than is ChatGPT (lingual). Those who extrapolate from current successes in AI to "AI will soon be able to do everything" fundamentally misunderstand AI.

However, full reliability is not always needed where AI *assists* rather than *replaces* humans - the next question.

--- How Do We Use AI Beneficially? (Q6)

Whether AI face recognition is beneficial or harmful depends, not just on the AI working properly or wrongly, but the role it plays and whether it is used with evil intent, carelessness or good intent. Nor will AI do all our jobs, as Elon Mush believes; similar predictions were made in the late 1970s!

Roles: Most popular discussion presupposes AI replacing humans, but AI can also assist humans. During the 1980s, I was involved in an AI system to advise managers on the strength of business sectors - analytical and economic aspect application. From information supplied by managers, it estimated sector strengths but then actively encouraged them to disbelieve it rather than accept its answers, inviting them to explore differences between their and its views. This revealed things they had overlooked, thus refining their knowledge. Knowledge refinement is the very opposite of AI replacing humans [FOOTNOTE 6].

Intent, at two levels: Whatever role, is AI used with good intent, evil intent or carelessness? Are decisions to invest in or deploy AI made with responsibility and wisdom, or with self-interest and fear of missing out?

--- Will AI Take Over From Humans? (Q2)

No. Because, to do so, it would have to (a) have encapsulated in its knowledge base the laws of every aspect (b) have done so completely and without errors or biases, including cultural. For the reasons discussed above I do not believe this is possible.

The danger from AI, in my opinion, is not AI capabilities but human sin. Humanity will tend to use AI in ways that are "affluent, arrogant and

unconcerned", which is the reason Sodom was destroyed and Judah was exiled [Ezekiel 16:49]. This attitude and mindset can affect all three human activities around the AI system, algorithm design, AI development, and AI use and deployment. Issues of climate change, biodiversity and the Global South are largely overlooked so far but, I submit, are more important issues in God's eyes, and for our future, than AI capability.

--- FOOTNOTES

1. = = = Link to Blog 1.

2. = = = Link to Blog 2.

3. [About aspects] These aspects were delineated in Dooyeweerd's philosophy, as described for example in "http://dooy.info/aspects.html". Other suites of aspects could be used, but Dooyeweerd's is most complete and most philosophically sound; see "http://dooy.info/compare.asp.html".

4. [About aspectual law, to clarify] Laws of the lingual aspect are deeper than laws of any given language, enabling language to occur - laws about linguistic syntax, semantics, pragmatics and so on.

5. [After modules of ChatGPT] For how ChatGPT works, see = = = = =

6. [about roles in which AI may be used] Basden [1983] outlines eight roles in which AI could be used and be beneficial. Strangely, there has been little discussion of roles since then, but most of the roles still apply today.
Basden A, (1983). On the application of Expert Systems. *Int. J. Man-Machine Studies*, 19:461-477. "http://kgsvr.net/andrew/-p/ai/Basden83-ApplicES.pdf"