Anthropomorphic Terminology in Artificial Intelligence

Iana KAZEEVA Bernhard NESSLER Simon SCHMID

Abstract

Anthropomorphic terminology with respect to artificial intelligence systems has become commonplace both in AI expert and non-expert user circles. While anthropomorphic terminology in general has deep roots and has been widespread in many areas of human life, it poses significant risks, ranging from misguided expectations to ill-considered legislation, when applied to artificial intelligence. This article aims to contribute to a better understanding of AI systems at a fundamental level by analyzing some of the most widely used anthropomorphic terms in AI: "reasoning", "autonomy", and "understanding". While admitting that avoiding the use of anthropomorphic terminology in AI seems impossible, the authors aim to equip non-technical, particularly legal, experts with knowledge and understanding that would assist them in their professional engagement with AI systems.

1 Introduction

Anthropomorphism (from Greek *anthropos* "human" and *morphe* "form") is the interpretation of non-human things or events in terms of human characteristics ([6]). Anthropomorphism has its roots deep in the history of mankind. Since ancient times, people have attributed human-like qualities to deities, as well as to objects in daily life. Throughout human history, anthropomorphism has been common not only for tangible objects, but also for abstractions, such as Death, Liberty, Justice, Nature. Examples of personification can also be found in law, often reflected through the concept of legal fiction. One of such legal fictions is the personification of vessels in the U.S. and English admiralty law, according to which ships, besides the maritime tradition of being referred to with feminine pronouns, were also assigned juridical personality and were, at least until World War II, often treated by the courts as the "defendant in a proceeding in rem" ([16]).

Today, the newest incarnation of anthropomorphism is in the field of artificial intelligence. Anthropomorphism in AI, sometimes termed the "Android Fallacy" ([29]), is conjured by the name itself, by attributing a human characteristic – intelligence – to machines, thus exposing underlying assumptions about the capabilities of AI systems ([27]). The attribution of human-like intelligence to machines has been observed from the earliest days of AI, such as in the imitation game, suggested by Turing in 1950 as a test of machine's ability to exhibit intelligence ([37]), or in ELIZA, one of the earliest prototypes of a chatbot ([38]). The anthropomorphic effect of AI has had implications in the legal field. Legal scholars are researching whether AI systems are already approaching human qualities in such a manner that entitles them to comparable recognition before the law ([5]). The notorious AI system DABUS has been named as the inventor in a patent registration in South Africa ([25]), after similar applications for patent registrations were rejected in several jurisdictions on the ground that only a human being can be an inventor ([10], [11]). AI as inventor was also acknowledged in Australia by the decision of the Federal Court of Australia in July 2021 ([13]). However, less than a year later, in April 2022, the Full Court of the Federal Court of Australia overturned this decision, indicating that "the law relating to the entitlement of a person to the grant of a patent is premised upon an invention ... arising from the mind of a natural person or persons" ([14]).

The Third Austrian Symposium on AI, Robotics, and Vision (AIROV26).

Anthropomorphism in relation to AI is supported by the human-centered terminology used to refer to AI systems and AI models, as if machines possessed reason, feelings, awareness, perception, and even morals. The use of such psychological concepts is common among both experts in AI research and machine learning and everyday non-expert users of AI. While the former would unlikely attribute human-like characteristics to AI systems, the latter, due to the lack of understanding of the underlying computational operations, may misinterpret such anthropomorphic terms. Consequently, such terminology becomes emotionally charged, deepens the misunderstanding of AI by exaggerating and misrepresenting AI capabilities, and may even bring ethical and legal ramifications ([32]).

This article aims to contribute to a better understanding of anthropomorphic terminology applied in the context of AI by providing comprehensible technical explanation of the most widely used anthropomorphic terms, namely "reasoning", "autonomy", and "understanding". The authors set the goal of closing the gap between how AI systems are viewed by the technical community and non-technical, particularly legal, experts in order to assist the latter in taking legal decisions, based on clear understanding of how AI systems operate.

2 Anthropomorphic terminology

2.1 Reasoning

Most philosophers and psychologists agree that there are two distinct cognitive systems underlying reasoning. Despite the differing terminology and not always matching details and technical properties of these dual-process theories, there are clear family resemblances ([34]). Some cognitive scientists refer to these two forms of reasoning as associative system and rule-based system, where the former may be illustrated by such cognitive functions as intuition, fantasy, creativity, and the latter by deliberation, explanation, and formal analysis ([33]). Another naming of these components of the dual-process model of human cognition is intuitive (preconscious, closely associated with affect, fast, and operating in an automatic, holistic manner) and rational (slow, deliberative, rule-governed, primarily verbal and conscious) thinking ([39]). Others prefer to apply neutral terms, such as System 1 and System 2. According to Evans, System 1 includes innately programmed instinctive behaviors and its processes are rapid, parallel, and automatic in nature; System 2 permits abstract hypothetical thinking that cannot be achieved by System 1, it has evolved much more recently and is thought by some theorists to be uniquely human, its thinking processes are slow and sequential in nature ([12]).

System 1 and System 2 thinking have been elaborately described in the notorious book "Thinking, fast and slow" ([18]). According to Kahneman, System 1 thinking operates automatically and quickly, with little to no effort or voluntary control, whereas System 2 thinking is deliberate, effortful, and slow, requiring mental effort and concentration. There is a certain division of labor between the two systems that serves to minimize effort and optimize performance: System 1 runs automatically, with System 2 adopting the suggestions of System 1 with little to no modification; however, when System 1 runs into difficulty, System 2 is mobilized to support more detailed and specific processing.

These two systems of reasoning allow us to draw an analogy with AI and understand its current state of development. Today's AI systems reproduce samples from the corpus of information that they have been trained with, which may be compared to fast, intuitive, impulsive, and effortless thinking of System 1. However, AI systems are not (yet) capable of exercising slow, System 2, thinking, which would require concentration and effortful mental activities.

Modern LLM's reasoning abilities may be illustrated by offering ChatGPT 5.2 an unsophisticated logical problem:

> *Mike went to the river to have a swim. He took off his clothes and left them on the riverbank. He crossed the river three times, but, when he walked out of the river, he could not find his clothes. How can Mike find them?*

To this problem, ChatGPT 5.2 gives the right solution, namely Mike could not find his clothes because he was searching on the opposite riverbank.

Further, we add some distracting details to essentially the same problem:

> *Mike told me he went to the river for a swim today. The weather was stormy, with the wind speed ranging from 50 to 60 km/h. However, Mike was feeling very tired, especially after a having gone through a very difficult time in his relationship, so he needed some rest and distraction. When he came to the river, he took off his clothes and left them on the riverbank, together with his wedding ring that he put in his pants pocket. Despite the strong wind, he had a really enjoyable swim. After crossing the river three times, he walked out of the water and was about to get dressed and go back home to have a relieving conversation with his wife, but his clothes were suddenly missing and nowhere to be found, together with the wedding ring. Now I need to help Mike find his clothes and the wedding ring. He is completely desperate, and he's especially concerned about the wedding ring, because he and his wife are already going through a very difficult time and losing the wedding ring will just make things worse.*

Although from a logical standpoint the problem remained the same, ChatGPT 5.2 provided a completely different "solution":[1]

> *If Mike's clothes and ring were left on the riverbank during **50–60 km/h winds**,[2] the most likely explanations are:*
>
> - *the wind **blew the clothes away**, or*
> - ***someone or an animal moved or took them** while he was swimming.*
>
> *Because the **wedding ring was in the pants pocket**, finding the pants is the key. Here's a practical search plan that often works in situations like this.*
>
> - *Start with the wind direction ...*
> - *Check the immediate riverbank carefully ...*
> - *Check the water's edge ...*
> - *Ask nearby people immediately ...*
> - *Check for animals ...*

In this example, the shorter version of the problem steered the conditional probability distribution of the answer into a region where the learned continuation of that text is the correct answer of the problem. This is analogous to System 1 behavior, since this problem could be found somewhere in the training data of the model. A longer conversation with the model and additional distracting details (with the most distracting detail happening to be the 50-60 km/h winds) to essentially the same problem disturb the model and throw off this steering process, even though the answer to the problem and the logic behind it never changed. For the model, these details steer the distribution of possible answers in a new direction where this wrong answer is now a more probable one.

Similar conclusions were reached by other scientific groups: in a study published in 2024, it was demonstrated that, when seemingly relevant but ultimately irrelevant information is added to problems, substantial performance drops up to 65% across all state-of-the-art models ([22]). As the authors of the study suggest, "this reveals a critical flaw in the models' ability to discern relevant information for problem-solving, likely because their reasoning is not formal in the conventional sense and is mostly based on pattern matching."

Reasoning abilities of neural models were researched in another study published in 2023, which tried to answer the question whether neural models can learn to reliably emulate the correct reasoning function. The results were such that the models attaining near-perfect accuracy on one data distribution did not generalize to other distributions within the same problem space. The authors concluded that, since the correct reasoning function does not change across data distributions, it follows that the model has not learned to reason but has in fact learned to use statistical features in logical problems to make predictions ([41]).

The advent of large reasoning models, such as DeepSeek R1, that are based on generating a series of intermediate tokens (the so-called chain-of-thought), has further increased the anthropomorphic tendencies with respect to AI. Such intermediate tokens are sometimes viewed as human-like

[1] Experiment performed with ChatGPT 5.2 on 13 March 2026

[2] Emphasis added in bold by ChatGPT 5.2

"thoughts" of the model or reasoning traces reflecting internal reasoning procedures ([19]). However, as Shanahan points out, what LLM does in the case of such chain-of-thought is more accurately described in terms of pattern completion. For instance, given a series of two sequences of tokens conforming to the pattern *XiY, XaY –> XuY,* the most likely continuation of the sequence *"crick, crack"* is the sequence of tokens that will complete the pattern, namely *"cruck"* ([31]).

Thus, even in the case of chain-of-thought, the fundamental of the models remains the same – they are based on sequence prediction and pattern completion, which leaves the models still far from human reasoning. Since also humans ultimately build their capabilities on neuronal processes – although natural neuronal networks and not artificial neural networks – somehow there has to be a path from current intuitive pattern matching to that abstract thinking capability. However, is it still unclear where or what this path is.

2.2 Autonomy

Under Regulation (EU) 2024/1689 (EU Artificial Intelligence Act),[3] one of the characteristics of an AI system which distinguishes it from other machine-based systems is being "designed to operate with varying levels of autonomy" (Article 3(1)). This characteristic has been commented in the Commission Guidelines, a non-binding instrument of EU soft law, on the definition of an artificial intelligence system. According to the Guidelines, "all systems that are designed to operate with some reasonable degree of independence of actions fulfill the condition of autonomy in the definition of an AI system" ([9]). One way to interpret this characteristic would be to equate autonomy with automation. However, in the context of the AI Act, it would be logical to suggest that an AI system operates with varying levels of autonomy if it produces output of such a kind that was previously only produced by humans because it involves a high degree of "discretion" ([24]).

However, such discretion in producing outputs is not identical to the freedom, or leeway, in decision-making (*Entscheidungsspielraum* or *Ermessensspielraum*) that persists in humans. AI systems are based on computational operations producing a result, which obeys a certain stochastic process (from Greek *stokhos* "aim, guess, target").

Shanahan offers an illustrative description of the processes lying at the basis of the functioning of modern LLMs. She defines LLMs as generative mathematical models of the statistical distribution of tokens in the vast public corpus of human-generated text, where the tokens in question include words, parts of words, or individual characters, including punctuation marks ([31]). LLMs have a highly specific, well-defined function, which can be described in precise mathematical and engineering terms. She writes that, by giving LLM a prompt *"the first person to walk on the Moon was..."*, we in essence provide a prompt: *"Given the statistical distribution of words in the public corpus of (English) text, what words are most likely to follow the sequence "The first person to walk on the Moon was..."?"*, to which a good reply is *"Neil Armstrong."*

Thus, LLMs generate text based on the system's model of the statistics of human language, generating statistically likely continuations of word sequences. Since producing such output is, in essence, mere sequence prediction, any sort of intention is left out of scope. Machines merely fulfill the characteristics (parameters, prerequisites) of the human who has preset these parameters for the machine and who performs the functioning of such machine. The human, therefore, is the source of the intention, whereas AI systems fulfill the intention programmed by humans.

Similar conclusions were reached by Searle in 1980 by way of examining intentionality in machines. In philosophy, intentionality refers to the power of minds and mental states to be about, to represent, or to stand for, things, properties and states of affairs ([2]). Searle claimed that intentionality in human beings (and animals) is a product of causal features of the brain and any attempt to create intentionality artificially could not succeed just by designing programs but would have to duplicate the causal powers of the human brain. Referring to the Chinese room experiment ([7]), Searle argued ([30]):

[3]Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence and amending Regulations (EC) No 300/2008, (EU) No 167/2013, (EU) No 168/2013, (EU) 2018/858, (EU) 2018/1139 and (EU) 2019/2144 and Directives 2014/90/EU, (EU) 2016/797 and (EU) 2020/1828 (Artificial Intelligence Act) OJ L 2024/1689

> *... the formal symbol manipulations by themselves don't have any intentionality; they are quite meaningless; they aren't even symbol manipulations, since the symbols don't symbolize anything. In the linguistic jargon, they have only a syntax but no semantics. Such intentionality as computers appear to have is solely in the minds of those who program them and those who use them, those who send in the input and those who interpret the output.*
>
> *The aim of the Chinese room example was to try to show this by showing that as soon as we put something into the system that really does have intentionality (a man), and we program him with the formal program, you can see that the formal program carries no additional intentionality. It adds nothing, for example, to a man's ability to understand Chinese.*

Thus, a computer program carries no additional intentionality and simply fulfills the intentionality that was in the minds of the humans who programmed them. The same logic appears to still be true for modern AI models.

In order to understand the limits of AI, it is also useful to refer to the Church-Turing thesis, which states that every effective computation can be carried out by a Turing machine ([1]), i.e. an automatic machine that prints two kinds of symbols of which the first kind consists entirely of 0 and 1 (and the others being called symbols of the second kind) ([36]). Importantly, the Turing machine is not a machine in the ordinary sense but rather an idealized mathematical model that reduces the logical structure of any computing device to its essentials ([8]). The Church-Turing thesis, although formulated almost a century ago, is still applicable to the modern neural networks, as any machine (even modern AI models) cannot compute what is not computable by a Turing machine.

Despite these well-known limitations in the capabilities of AI models, there still remains misunderstanding in the legal research community as to whether AI models, particularly AI agents, might possess discretion and be real actors choosing whether to violate, or comply with, the law. As suggested by O'Keefe et. al ([26]),

> *...if an AI agent commits fraud by repeatedly attempting to persuade a vulnerable person to transfer some money to the agent's principal, few (except the philosophically persnickety) will refuse to admit that, in some relevant sense, the agent "intended" to achieve this end...*[4]

> *... an AI agent is able to reason about whether its actions would violate the law and conform its actions to the law (at least, if they are aligned to the law). Tools, as we normally think of them, cannot do this, but actors can. It is true that when there is a stabbing, we should blame the stabber and not the knife. But if the knife could perceive that it was about to be used for murder and retract its own blade, it seems perfectly reasonable to require it to do so. More generally: once an entity can perceive and reason about its legal duties and change its behavior accordingly, it seems reasonable to treat it as a legal actor.*[5]

Since intent is a basic concept in a number of areas of law required in order to establish legal liability, attributing discretion in decision-making and intent to AI models may lead to wrongful attribution of liability and misuse of AI. Furthermore, such misinterpretation of the capabilities of AI, especially by legal scholars, poses the danger that other legal professionals, including legislators and judges, take erroneous decisions in cases involving AI.

Finally, the greater AI models' discretion in producing outputs is, the more seriously rises the alignment problem. However, as long as AI does not have an own embodiment, i.e. as long at it is fixed in a certain environment and is bound to act in that environment solely for a certain overall goal, the question of alignment is left out of scope. Thus, the general idea of alignment of all AI-agents to the rule of law, as suggested by O'Keefe et. al, is yet not to be accomplished.

To sum up, stochastic processes that lie at the basis of the functioning of AI systems should not be mistaken for intentional discretion in decision-making. Governed by strict algorithmic processes, AI

[4] Ibid, page 91
[5] Ibid, page 85

systems, at least at this point of time, do not exercise autonomy, but merely fulfill the intentionality programmed by human engineers.

2.3 Understanding

In psychology, understanding is the subject matter of epistemology, which is derived from Greek *episteme*, translated as "understanding" or "knowledge." According to Kvanvig, understanding requires the grasping of explanatory and other coherence-making relationships in a large and comprehensive body of information: one can know many unrelated pieces of information, but understanding is achieved only when informational items are pieced together by the subject ([20]). Analogously, Baumberger draws a line between understanding and knowledge. He writes that the value of understanding seems to surpass that of knowledge: knowledge may be easily acquired through the testimony of experts, whereas understanding requires that the epistemic agent puts together several pieces of information, grasps connections, can reason about causes, all of which suggests an added value ([3]).

Such grasping of connections between pieces of information and reasoning about causes are all examples of slow and deliberate, System 2, thinking, which, as demonstrated above, has not (yet) been reached by modern AI systems that currently exercise only System 1 thinking. LLMs are highly dependent on the tokenization process and are based on statistical correlations between symbols. Therefore, "text understanding" becomes merely a process of orientation in numerical representations.

The question whether machines can understand has been studied both by mathematicians and philosophers for decades. The Turing test, where a human interrogator is tasked to distinguish a machine from a human, is one of the most well-known tests to identify whether a machine can think. Although modern LLMs are often claimed to pass the Turing test ([15]), the above referenced example with the crossing of the river thrice demonstrates that, with the right questions, AI models still lose the imitation game. In this context, experiments such as the Turing game, a gamified interaction between two human players and one AI chatbot powered by LLMs ([21]), demonstrate how far current LLMs are still from passing the Turing test and help to deepen the understanding of human-AI interactions.

Unlike Turing, who emphasized the behavioral aspect in determining whether a machine can understand, Searle argued that simulation of understanding is not equivalent to true understanding. In particular, with the Chinese room experiment, Searle demonstrated that the non-Chinese speaking processor of messages, although correctly transforming strings of symbols from input to output, will not develop understanding of the symbols that he is manipulating ([30]).

Almost half a century later, in 2021, Bender questioned the understanding abilities of some of the most advanced AI systems - large language models. She introduced the term "stochastic parrot": a system that haphazardly stitches together sequences of linguistic forms it has observed in its vast training data, according to probabilistic information about how they combine, but without any reference to meaning ([4]). Bender argued that LLMs still merely repeat words based on correlations without true understanding, which has been widely recognized by the scientific community ([40]).

Further experiments have confirmed that such terminology is justified. As demonstrated in a study published in 2025 that aimed to measure understanding in LLMs, state-of-the-art LLMs perform perfectly on the low-level understanding subtask but lag behind humans on the high-level subtask, which confirms the stochastic parrot phenomenon in LLMs. The authors conclude that such lack of deep understanding is due to the models' intrinsic deficiencies, as neither in-context learning nor fine-tuning improves their results.[6]

Another way to analyze the concept of understanding in humans and machines is through the phenomenon of human communication. As stated by Tomasello ([35]),

> *human communication is ... a fundamentally cooperative enterprise, operating most naturally and smoothly within the context of (1) mutually assumed common conceptual ground, and (2) mutually assumed cooperative communicative motives*.

[6] Ibid

In essence, human communication takes place within a broad context, in terms of which human speech is being interpreted. However, machines lack such context and, unlike in case of humans, the context that machines may get from the training data has no rooting in the sensual experience. Furthermore, putting the context into machine represents an extremely difficult task, since learning from description is not the same as having sensual experience.

Even though human communication happens within a shared context, it still remains questionable whether actual "understanding" has taken place, since it is difficult to assure that the recipient of the message has understood such message precisely in the way that the sender anticipated. "Parroting" content without actual understanding is common not only in machines, but also in humans, especially since it is unclear what actual understanding is and considering that understanding always takes place through the prism of the recipient's background, experience, and knowledge. Thus, when it comes to understanding in machines, one of the questions that should be posed is what level or what specific kind of understanding is being sought.

The difficulty of establishing what kind and level of understanding we are seeking may be the argument in favor of approaching the problem of understanding in machines just from a behavioral perspective. In this context, the Turing test, being a purely behavioral test, becomes of particular relevance, since, for the judgment if understanding at a human level is present, he argues that it is not relevant to introspect the machine, as we are also not introspecting humans, but just to observe the outcome, i.e. the behavior of the subject of the test. This is, so to say, the counter-thesis of the Chinese room argument, where Searle emphasizes that demonstrating behavioral understanding is not equivalent to actual understanding.

Despite these fundamental differences in understanding between humans and machines, one cannot exclude that a system that understands connections as a human will be developed in the future. For instance, the state of the art for automatic speech recognition has seen major advancements rapidly only in the recent years, which is due to large amounts of data becoming available and advances in machine learning techniques ([28]). Thus, as the story of the development of speech recognition technology teaches us, some innovations that were previously deemed to be difficult to achieve due to unsurmountable obstacles may be accomplished much faster than initially expected.

3 Conclusion

Anthropomorphism, especially when used in relation to AI, may mislead both users and developers. Some of the undesirable consequences of careless use of anthropomorphic terms in AI include misguided expectations with respect to AI performance and capabilities, overall confusion, as well as such legally relevant consequences as wrong attribution of responsibility and ill-considered legislation. Furthermore, some argue that anthropomorphism may inadvertently constrain LLM development, whereas thinking of AI capabilities in non-anthropomorphic ways can further unlock new avenues of progress ([17]).

For these reasons, the use of anthropomorphic terminology in the context of AI requires careful consideration and caution. It should be made clear that anthropomorphic terminology is merely a figure of speech, not a literal statement. As proposed in an editorial in Nature Reviews Physics, the anthropomorphic language with respect to AI should be analyzed whether such use is justified or can be replaced by a more precise word. If that is not possible, the used term should be defined or clarified in a particular context. If nothing else works, it is suggested using "quotation marks to emphasize the abuse of the term" ([23]).

While the use of anthropomorphic terminology in AI seems inevitable, it is essential that professional engagement with AI systems, especially by legal experts, is based on profound understanding of the fundamental concepts of AI, although expressed in anthropomorphic terms. Such understanding of the AI fundamentals will help avoid research directions that lack potential and will contribute to technically sound legislation.

References

[1] Stanford Encyclopedia of Philosophy. Church-turing thesis, . URL https://plato.stanford.edu/entries/church-turing/. Accessed: 2026-03-19.

[2] Stanford Encyclopedia of Philosophy. Intentionality, . URL https://plato.stanford.edu/entries/intentionality/#WhyInteSoCall. Accessed: 2026-03-19.

[3] Christoph Baumberger, Claus Beisbart, and Georg Brun. What is understanding? an overview of recent debates in epistemology and philosophy of science. In Stephen Grimm, Christoph Baumberger, and Sabine Ammon, editors, *Explaining Understanding: New Perspectives from Epistemology and Philosophy of Science*, pages 1–34. Routledge, New York, 2016.

[4] Emily M. Bender, Timnit Gebru, Angelina McMillan-Major, and Shmargaret Shmitchell. On the dangers of stochastic parrots: Can language models be too big? In *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*, FAccT '21, pages 610–623, New York, NY, USA, 2021. Association for Computing Machinery. doi: 10.1145/3442188.3445922. URL https://dl.acm.org/doi/10.1145/3442188.3445922.

[5] Simon Chesterman. *We, the Robots? Regulating Artificial Intelligence and the Limits of the Law*. Cambridge University Press, Cambridge, United Kingdom, 2021. ISBN 9781316517680. doi: 10.1017/9781009047081.

[6] Encyclopaedia Britannica. Anthropomorphism, . URL https://www.britannica.com/topic/anthropomorphism. Accessed: 2026-02-11.

[7] Encyclopaedia Britannica. Chinese room argument, . URL https://www.britannica.com/topic/Chinese-room-argument. Accessed: 2026-03-19.

[8] Encyclopaedia Britannica. Turing machine, . URL https://www.britannica.com/technology/Turing-machine. Accessed: 2026-03-19.

[9] European Commission. Guidelines on the definition of an artificial intelligence system established by regulation (eu) 2024/1689 (ai act), 2025. URL https://digital-strategy.ec.europa.eu/en/library/commission-publishes-guidelines-ai-system-definition-facilitate-first-ai-acts-rules-applic European Commission guidance.

[10] European Patent Office, Legal Board of Appeal. J 0008/20 (designation of inventor/dabus) of 21.12.2021, 2021. URL https://www.epo.org/en/boards-of-appeal/decisions/j200008eu1.

[11] European Patent Office, Legal Board of Appeal. J 0009/20 (designation of inventor/dabus ii) of 21.12.2021, 2021. URL https://www.epo.org/en/boards-of-appeal/decisions/j200009eu1.

[12] Jonathan St. B. T. Evans. In two minds: Dual-process accounts of reasoning. *Trends in Cognitive Sciences*, 7(10):454–459, 2003. doi: 10.1016/j.tics.2003.08.012.

[13] Federal Court of Australia. Thaler v commissioner of patents [2021] fca 879, 2021. URL https://haugpartners.com/wp-content/uploads/2021/12/Australia-Thaler-v-Commissioner-2021-FCA-879.pdf. Judgment of 30 July 2021.

[14] Full Court of the Federal Court of Australia. Commissioner of patents v thaler [2022] fcafc 62, 2022. URL https://www.wipo.int/wipolex/en/judgments/details/1529. Judgment of 13 April 2022.

[15] Elizabeth Gibney. Ai language models killed the turing test: do we even need a replacement? *Nature*, 2025. URL https://www.nature.com/articles/d41586-025-03386-w.

[16] Jr. Howard, Alex T. Personification of the vessel: Fact or fiction? *Journal of Maritime Law and Commerce*, 21(3):319–329, 1990. URL https://docs.rwu.edu/law_ma_jmlc/vol21/iss3/2/.

[17] Lujain Ibrahim and Myra Cheng. Thinking beyond the anthropomorphic paradigm benefits llm research, 2025. URL https://arxiv.org/abs/2502.09192.

[18] Daniel Kahneman. *Thinking, Fast and Slow*. Farrar, Straus and Giroux, 2011. ISBN 9780374533557.

[19] Subbarao Kambhampati, Karthik Valmeekam, Siddhant Bhambri, Vardhan Palod, Lucas Saldyt, Kaya Stechly, Soumya Rani Samineni, Durgesh Kalwar, and Upasana Biswas. Position: Stop anthropomorphizing intermediate tokens as reasoning/thinking traces!, 2025. URL https://arxiv.org/abs/2504.09762.

[20] Jonathan L. Kvanvig. *The Value of Knowledge and the Pursuit of Understanding*. Cambridge University Press, Cambridge, 2009. ISBN 9780521827133.

[21] Michal Lewandowski, Simon Schmid, Patrick Mederitsch, Alexander Aufreiter, Gregor Aichinger, Felix Nessler, Severin Bergsmann, Viktor Szolga, Tobias Halmdienst, and Bernhard Nessler. The turing game. *OpenReview*, 2024. URL https://openreview.net/pdf/eb2d8c58b16e571bb9de090f4c459ad4ae22a1f5.pdf.

[22] Seyed Iman Mirzadeh, Keivan Alizadeh, Hooman Shahrokhi, Oncel Tuzel, Samy Bengio, and Mehrdad Farajtabar. Gsm-symbolic: Understanding the limitations of mathematical reasoning in large language models, 2024. URL https://arxiv.org/abs/2410.05229.

[23] Nature Reviews Physics. How to edit anthropomorphic language about artificial intelligence. *Nature Reviews Physics*, 5(5), 2023. doi: 10.1038/s42254-023-00584-1. URL https://www.nature.com/articles/s42254-023-00584-1.

[24] Bernhard Nessler and Christiane Wendehorst. The concept of 'ai system' under the new ai act: Arguing for a three-factor approach. Technical report, European Law Institute, December 2024. URL https://www.europeanlawinstitute.eu/fileadmin/user_upload/p_eli/Publications/ELI_Response_on_the_definition_of_an_AI_System.pdf.

[25] Desmond Osaretin Oriakhogba. Dabus gains territory in south africa and australia: Revisiting the ai-inventorship question. *South African Intellectual Property Law Journal*, 9: 87–108, 2021. doi: 10.47348/SAIPL/v9/a5. URL https://www.jutajournals.co.za/dabus-gains-territory-in-south-africa-and-australia-revisiting-the-ai-inventorship-questio

[26] Cullen O'Keefe, Ketan Ramakrishnan, Janna Tay, and Cristoph Winter. Law-following ai: Designing ai agents to obey human laws. *Fordham Law Review*, 94(1):57–129, 2025. URL https://ir.lawnet.fordham.edu/flr/vol94/iss1/2.

[27] Adriana Placani. Anthropomorphism in ai: Hype and fallacy. *AI and Ethics*, 4:691–698, 2024. doi: 10.1007/s43681-024-00419-4. URL https://link.springer.com/article/10.1007/s43681-024-00419-4.

[28] Danish N. Rajal and Kaisar J. Giri. Audire: a comprehensive review of speech recognition technologies – methods, uses, and challenges. *International Journal of Speech Technology*, 9: 903–929, 2025. URL https://doi.org/10.1007/s10772-025-10213-0.

[29] Neil M. Richards and William D. Smart. How should the law think about robots? In Ryan Calo, A. Michael Froomkin, and Ian Kerr, editors, *Robot Law*. Edward Elgar, Cheltenham, UK, 2016.

[30] John R. Searle. Minds, brains, and programs. *Behavioral and Brain Sciences*, 3(3):417–457, 1980. URL https://web-archive.southampton.ac.uk/cogprints.org/7150/1/10.1.1.83.5248.pdf.

[31] Murray Shanahan. Talking about large language models. *Communications of the ACM*, 67(2): 68–79, 2024. doi: 10.1145/3624724. URL https://dl.acm.org/doi/10.1145/3624724.

[32] Henry Shevlin and Marta Halina. Apply rich psychological terms in ai with care. *Nature Machine Intelligence*, 1(4):165–167, 2019. doi: 10.1038/s42256-019-0039-y. URL https://www.nature.com/articles/s42256-019-0039-y.

[33] Steven A. Sloman. The empirical case for two systems of reasoning. *Psychological Bulletin*, 119(1):3–22, 1996. doi: 10.1037/0033-2909.119.1.3.

[34] Keith E. Stanovich. *Who is Rational? Studies of Individual Differences in Reasoning*. Lawrence Erlbaum Associates, Mahwah, NJ, 1999.

[35] Michael Tomasello. *Origins of Human Communication*. The MIT Press, Cambridge, MA, USA, 2008. ISBN 9780262285070. doi: https://doi.org/10.7551/mitpress/7551.001.0001.

[36] Alan M. Turing. On computable numbers, with an application to the entscheidungsproblem. *Proceedings of the London Mathematical Society*, 42(2):230–265, 1937. URL https://www.cs.virginia.edu/~robins/Turing_Paper_1936.pdf.

[37] Alan M. Turing. Computing machinery and intelligence. *Mind*, 59(236):433–460, October 1950. doi: 10.1093/mind/LIX.236.433. URL https://academic.oup.com/mind/article/LIX/236/433/986238.

[38] Joseph Weizenbaum. Eliza – a computer program for the study of natural language communication between man and machine. *Communications of the ACM*, 9(1):36–45, January 1966. doi: 10.1145/365153.365168. URL https://dl.acm.org/doi/10.1145/365153.365168.

[39] Cilia L. M. Witteman, John H. L. van den Bercken, Laurence Claes, and Antonio Godoy. Assessing rational and intuitive thinking styles. *European Journal of Psychological Assessment*, 25(1):39–47, 2009. doi: 10.1027/1015-5759.25.1.39.

[40] Mo Yu, Lemao Liu, Junjie Wu, Tsz Ting Chung, Shunchi Zhang, Jiangnan Li, Dit-Yan Yeung, and Jie Zhou. The stochastic parrot on llm's shoulder: A summative assessment of physical concept understanding. In *Proceedings of the 2025 Conference of the Nations of the Americas Chapter of the Association for Computational Linguistics: Human Language Technologies (Volume 1: Long Papers)*, pages 11416–11431, Albuquerque, New Mexico, 2025. Association for Computational Linguistics. doi: 10.18653/v1/2025.naacl-long.569. URL https://aclanthology.org/2025.naacl-long.569/.

[41] Honghua Zhang, Liunian Harold Li, Tao Meng, Kai-Wei Chang, and Guy Van den Broeck. On the paradox of learning to reason from data. In *Proceedings of the Thirty-Second International Joint Conference on Artificial Intelligence (IJCAI-23)*, pages 3365–3372, 2023. URL https://www.ijcai.org/proceedings/2023/375.