

# Report on the GChACM Symposium Why we should abandon the term Artificial Intelligence

---



Thomas Matzner

GChACM Council Member for Computing Machinery & Ethics

<https://informatikethik.de>

Am 16. Mai 2025 führte das German Chapter of the ACM dieses Symposium in Form einer Videokonferenz durch. Die Grundlage war ein Positionspapier von Thomas Matzner, auf das in zwei Positionspapieren von Martin Grundmann und Tobias Hamberger geantwortet wurde. Während des Symposiums fand nicht nur eine inspirierte und inspirierende Diskussion zu dem Thema und verbreiteten Missverständnissen und Irrtümern statt, mit denen es verbunden wird. Die Teilnehmer vereinbarten auch, eine gemeinsame Stellungnahme zu versuchen, um solchen Irrtümern zu begegnen und innerhalb wie außerhalb der Fachwelt für Klarheit über den Begriff der Künstlichen Intelligenz und berechnete Erwartungen daran zu schaffen.

Nachtrag zu den Positionspapieren: Soeben ist, herausgegeben von Martin Hähnel und Regina Müller, „A Companion to Applied Philosophy of AI“ erschienen, in dessen zweitem Beitrag von Vincent C. Müller „Philosophy of AI: A Structured Overview“ Bestrebungen erwähnt werden, den Begriff durch andere zu ersetzen.

Die Positionspapiere sind im Folgenden wiedergegeben. Wir freuen uns über Kommentare und Beiträge.

## Position Paper Thomas Matzner

January 2025

The subject Artificial Intelligence (AI) comprises a collection of mathematical problem solving methods. Some of them are concerned with logical reasoning, others employ stochastic methods to obtain results in fields where reliable rules for such reasoning are missing, e.g. predicting human behavior. Those methods are widely taught in courses for computing professionals, there exists a solid stock of literature used for teaching them, and a huge field of research about them.

The term Artificial Intelligence is also used in a different sense. The anthropomorphic term „intelligence“ seduces to compare the behavior of machines running AI techniques to human behavior. Even the textbooks mentioned above, in their introductory chapters, raise speculations about machines being nearly as intelligent, as intelligent, or even more intelligent than humans, in one, many, or „all“ fields. I call them speculations because intelligence isn't precisely defined even when applied to humans, and the idea what a machine has to achieve to be called intelligent is even more hazy. One of the very first computers, used by Alan Turing during world war 2 to decode the wireless messages of the German navy, solved this task much faster and more reliable than humans might have done and might therefore be called intelligent. But those early machines and the software running on them are generally not included in the AI notion we discuss today.

The computing pioneer and ACM Turing Award Winner Edsger W. Dijkstra gave this warning: (Dijkstra 1975):

*The use of anthropomorphic terminology when dealing with computing systems is a symptom of professional immaturity.*

While in those times this was mainly an admonition towards the computing community, nowadays the anthropomorphization of machines is widely applied and predominates the public reception of AI, compared with a down-to-earth assessment what it really is. A few popular ramifications of this are:

- The (techno) responsibility gap (Friedman 2024) that is claimed to arise when an algorithm is run on a machine without human supervision (which has regularly been done during the past 80 years). Authors claim that the operator responsible for this hands the responsibility away from himself. Opinions differ whether the responsibility just vanishes or is transferred to the machine. Of course, no such gap actually exists, but the human actor stays responsible for all his actions, no matter whether he is present when they occur. We should not let his responsibility vanish, but should even increase it, if he fails to understand or control them.
- The paperclip apocalypse, a thought experiment concerning a super intelligent machine designed for making paperclips in the most efficient manner. It is suspected that this machine might exaggerate its actions, trying to turn large parts of the world into paperclips. When humans tried to switch it off, it would consider them its enemies and start fighting them, probably also turning them into paperclips. The philosopher who first published this idea didn't notice that such a machine could hardly be called super intelligent.
- In a recently published handbook of the ethics of AI, a theologian discusses the relationship between theological, scientific, technological, and finally AI ideas (Smith 2024). Among other things, he refers to the story of God creating the world, including humans, in the book Genesis of the Bible. He doesn't notice the obvious difference: Christians praise God for his

acts of creation, but they don't hold him responsible for the outcome in the sense of criticising him for the deficiencies of his creatures, in particular those of humans. In contrast, humans creating an artifact can and shall be made responsible for its behavior, which makes it their duty to formulate an explicit purpose, to follow design principles, and to make sure that the result fulfils the purpose without causing unwanted collateral damage. Nothing of this may be learned from Genesis.

- Numerous stories get told about AI systems that start getting their own ideas about their operation purpose and turn against the people they are supposed to assist. A prototype of this kind of story is the German film HAL from 2016 in the Tatort TV series (HAL 2016). These stories have in common that they hardly convey any insights about AI systems but are products of science fiction.

Instead of attempting to describe the mathematical methods mentioned in the beginning using the hazy term „intelligence“, I suggest a perception that shows us more clearly what these techniques are, what is remarkable about them and how we can use them to provide value and avoid damage. Many of these techniques are stochastic and therefore offer themselves to be applied for operation purposes which are not governed by proven logical rules. Scarcely anyone would try to use stochastic methods for designing an electronic circuit, for accounting or tax calculation. Other operation purposes, however, aren't determined by proven rules like those, many of them because they involve the prediction of the behavior of single humans, or groups, or masses of those, including economic development. As no proven calculus for human behavior exists, people have invented other methods to come to grips with such operation purposes. Among them are approximate rules, like „a student having good grades in mathematics will become a good software developer“, or „an offender with three previous convictions is more likely to reoffend than one without any“. Among them also is human intuition.

**Operation purposes governed by stochastics:** For those, new solution paths have opened with the availability of powerful computation, data transmission, and data storage mechanisms at reasonable cost. This is what has driven the current AI hype. But the question which solution path to take is not the first and foremost one. Asking for „more AI usage“ or, conversely, for its rigid control, is similar to asking for using more psychology or less electricity. The core question is the admissibility of the operation purpose.

- Before asking whether politicians should be allowed to use generative AI during their election campaigns, we should ask whether campaigning as such, be it for political convictions, industry products, or cultural creations, is admissible and where the limits should be set.
- Before asking whether AI may be used to predict whether a student will succeed in a discipline, we should ask what we expect from our educational system: to give everyone the chance to learn what she or he desires to, or to lead students toward those disciplines in which they promise to make the best contribution for common welfare.

Answering such moral questions is often inherently difficult, as the science of ethics doesn't offer us undisputable methods for this task. A simple argument illuminating this difficulty is David Hume's principle, stating that it is logically impossible to deduct moral rules from empirical facts, where only the latter are scientifically provable. Once, however, a certain operation purpose has been identified and agreed on as permissible, it's comparatively easy to identify the solution path or paths to achieve

it: The best solution path, the one that offers the best chance to actually achieve the operation purpose, is to be preferred.

Finding the best, or at least a good, solution path may require time and effort. The safest way is, for each operation purpose separately, to conduct scientific research that compares the effectiveness of a large number of solution paths. Traditionally, this step is omitted for the vast majority of operation purposes. Politicians justify their convictions using easy-to-swallow phrases. People deciding about admission of candidates for studying a discipline or performing a job use an individual mixture of approximate rules, personal convictions, and intuition. Most laws allow judges a large range for assessing the guilt and prospects of offenders. Mankind has got used since hundreds and thousands of years to teachers assessing students, superiors assessing workers, and judges assessing offenders and witnesses.

How do stochastic techniques come into this game? First, they must be bound to each imaginable concrete operation purpose. One and the same technique may be unacceptable for operation purpose A, because the operation purpose itself is unacceptable, may be a bad choice for operation purpose B, because there are better solution paths for it, and may be ideal for operation purpose C. (Kahneman 2011, 222..233) quotes extensive research proving that, for predictions under uncertainty, simple approximate rules perform equally or better than human intuition, so the belief that humans are superior to formal methods is at least to be questioned in every such case. To get a clear picture of the quality of solution paths, it will be necessary, for each admissible operation purpose, to compare the actual effectiveness of the commonly used approximate rules, of several stochastic methods, and of previous decisions made by humans.

This sounds cumbersome and costly, and in many cases it really is. But there are a few operation purposes in which broad scientific evaluation of solution paths gets done, no matter how costly it is. One of them is the development and admission of medical treatment, including pharma products. Here, a scientific proof of effectiveness is required, and the time and effort of conducting that proof constitutes most of the high price of many of those products. Another field is fully automatic driving of cars, where tests have been made since many years ago and will probably continue to be made until – possibly, not certainly yet – such cars will be offered to consumers. For other operation purposes, making sure they are reliably fulfilled might be equally costly. But we should be grateful that the advent of stochastic algorithmic techniques has drawn our attention to performing such tests. As long as human behavior was predicted by other humans, possibly using approximate rules, no such tests were done, and we learned to live with the uncertainty of each decision maker using their own rules and prejudices.

So, the task to solve is to identify those operation purposes for which human societies find it worthwhile to identify good solution paths. Research on those needs to clarify which type of solution path is optimal or at least good enough. Stochastic methods are just candidates in this challenge and need to be soberly compared to their alternatives. Anthropomorphization is neither needed nor beneficial for this task.

And what remains of the term AI?

*„Artificial Intelligence“ is a historic term denoting a collection of mathematical problem solving techniques that, while being explored since the 1950s, required, for most purposes, more machine resources than were available at that time.*

## Acknowledgements

Most of the thoughts included here came to life during a stay in the wonderful holiday residence of Denise Berger and Ulrich Koppermann who invited me to join them there, close to the Mediterranean beach.

Thanks go to Wolfgang Keller and Marc-Oliver Pahl for reading and commenting on a draft of this paper, which led to tightening quite a few passages.

## References

Dijkstra, Edsger Wybe. *How do we tell truths that might hurt?* 18. 06 1975.

<https://www.cs.virginia.edu/~evans/cs655/readings/ewd498.html> (accessed 19. 08 2024).

Friedman, Cindy. „Responding to the (techno) Responsibility Gap(s).“ In *Handbook on the Ethics of Artificial Intelligence*, editor: David J. Gunkel, 326. Cheltenham: Edward Elgar Publishing, 2024.

HAL. <https://www.daserste.de/unterhaltung/krimi/tatort/sendung/hal-104.html>. 2016.

Kahneman, Daniel. *Thinking, Fast and Slow*. London: Penguin, 2011.

Smith, Joshua. „Faith, Technology, and the Ethics of AI.“ In *Handbook on the Ethics of Artificial Intelligence*, editor: David J. Gunkel. Cheltenham: Edward Elgar Publishing, 2024.

# Positionspapier (Martin Grundmann)

Eine Verbannung des Begriffes Artificial Intelligence bzw. Künstliche Intelligenz (AI/KI) aus der nicht-fachlichen Öffentlichkeit wird nicht möglich sein. Er ist schon zu lange als (undefiniertes) Konzept, Marketing-Begriff und als Buzzword „in aller Munde“.

Der Begriff AI hat sich in der nicht-fachlichen Öffentlichkeit spätestens von den zugrunde liegenden Algorithmen gelöst. Diese Öffentlichkeit versteht Algorithmus häufig sogar als Synonym für AI. Dadurch entstand ein breites Verständnis, dass AI eine andere, eingeschränkte Form von Intelligenz ist.

Spätestens mit der extremen Skalierung kombinierter Algorithmen für die LLMs sind die Algorithmen praktisch nicht mehr identifizierbar – daher die immer wiederkehrenden Forderungen nach Erklärbarkeit und Interpretierbarkeit. Letztere werden einerseits auch wieder durch Algorithmen geschaffen, sind andererseits aber auch ein weiterer Schritt zur Intelligenz. Andererseits laufen auch in unserem Körper nicht nur in Gehirn und Nervensystem hochgradig skalierte, kombinierte physikalische und chemische Prozesse ab. Selbst wenn wir die einzelnen Prozesse verstehen und mit Algorithmen beschreiben könnten, würden wir vermutlich trotzdem nicht verstehen, wie das System menschliche Intelligenz bewirkt. Gibt es überhaupt eine breit akzeptierte Definition der „menschlichen Intelligenz“?

Das Wort Intelligenz halte ich persönlich für nicht ganz falsch. Einerseits haben die AI-Systeme durchaus Eigenschaften, die man früher als intelligent bezeichnet hätte. Andererseits gibt es jetzt schon einen großen Druck, die AI als Verstärker der menschlichen Intelligenz einzusetzen, aber auch als deren Ersatz. Um das in vernünftige (intelligente ?) Bahnen zu lenken, werden wir ohnehin über Intelligenz reden müssen – zunächst die menschliche und dann die Ergänzung durch die maschinelle.

Wissenschaft und Fachwelt sollten es besser machen als die Öffentlichkeit, z.B. durch einen bewussten Umgang mit den Begriffen AI/KI sowie Intelligenz. Zusammen mit der GI oder dem ACM könnten wir vielleicht so eine Diskussion in größerem Rahmen initiieren. Mindestens sollte man für Vorträge, Fachartikel, Meetings, Dokumente, ... definieren, wie darin Intelligenz bzw. AI/KI verstanden werden soll.

Wie immer wir bzgl. des Begriffes AI und dem Einsatz von AI „intelligent“ weitermachen wollen: **Wir brauchen unbedingt ein halbwegs gemeinsames Verständnis, wie wir Intelligenz im Allgemeinen und menschliche/künstliche Intelligenz definieren.**

Neben der Definition von „Intelligenz“ finde ich es sehr wichtig zu „philosophieren“, **wo und warum Intelligenz für die Welt, die Menschheit, oder unsere Gesellschaft wichtig ist.** Wo und wie könnten die Menschen dabei Unterstützung von Maschinen und ihre Algorithmen brauchen, um noch besseren Nutzen aus der Intelligenz zu ziehen? Welche Nebenwirkungen und Risiken einer maschinellen Intelligenz können wir heute schon identifizieren?

Beispiel: Vielleicht könnten wir uns darauf einigen, dass flexible Durchdringung von unbekanntem, komplexen Systemen und die Abstraktion eine intelligente Eigenschaft ist. Hier gibt es Grenzen, wieviel ein (durchschnittlicher) Mensch intellektuell leisten kann. Es wird immer aufwändiger, diese Eigenschaft zu skalieren. Mit Unterstützung durch „AI“ ist heute in einigen Bereichen eine weitere Skalierung möglich geworden. Diese Skalierung könnte allerdings mit ihrer höheren Produktivität die Komplexität und das Chaos auch vergrößern, z.B. durch intelligente fakes, durch kriminellen Einsatz, oder durch unintelligenten Ersatz menschlicher Intelligenz.

## Symposium "Why we should abandon the term Artificial Intelligence"

*Diskussionsbeitrag von Tobias Hamberger*

Why we should abandon the term „Artificial Intelligence“?

Auf diese Frage gab der Eröffnungsbeitrag bereits eine Antwort. Dort wurde unter anderem kritisiert, dass der Ausdruck „Artificial Intelligence“ falsche Vorstellungen und falsche Erwartungen zu diesem Thema (mit technischen und ethischen Implikationen) verursacht.

In meinem Beitrag beleuchte ich speziell den Begriff „AI Ethics“ genauer, der auch etwas unscharf sein kann, weil verschiedene Akteure auf dem Feld der „AI“ sehr unterschiedliche Dinge darunter verstehen. Deshalb sollte auch dieser Begriff kritisch betrachtet werden.

Denn haben Beteiligte ein völlig unterschiedliches Verständnis von KI-Ethik, ohne sich dessen bewusst zu sein, kann dies meiner Ansicht nach zu Enttäuschung und gegenseitiger Entfremdung führen, obwohl stattdessen eher Zusammenarbeit und gegenseitige Ergänzung von ethischen Sichten gefragt wäre.

Neben einschlägigen Beispielen, in denen der Begriff „AI Ethics“ zu „Whitewashing“-Zwecken missbraucht wird, wären Beispiele:

- Aufgrund der ökologischen und gesellschaftlichen Folgen des Einsatzes von KI erwarten viele von einer KI-Ethik, dass sie wirksam diese Folgen eindämmt. Im Rahmen von ethischen Betrachtungen können sich dann zwar Prinzipien wie „Fairness“ oder „Freedom from Bias“ ergeben, die eine KI-Anwendung erfüllen muss, jedoch verliert sich auf der einen Seite die Umsetzung dann oft in technische Details und das eigentliche gesellschaftliche Problem dahinter wird aus den Augen verloren. Auf der anderen Seite bleiben die (evtl. überzogenen) Erwartungen der Kritikerseite unerfüllt und verstärkt deren Ablehnung von „KI“.
- Es ist lobenswert, dass viele große Firmen bereits Policies und Prinzipien für den Umgang mit KI erstellt haben, die auf dem Papier durchaus einen soliden Eindruck machen. Das sorgt einerseits für einen großen Vertrauensvorschuss in der Öffentlichkeit. Andererseits ist aber auch zu beobachten, dass von Firmen selbst erstellte Policies nicht verbindlich sind, und durch eigenmächtige Einschränkungen dieser Vertrauensvorschuss wieder verspielt werden kann.

Durch unterschiedliche Interpretation des Begriffs „AI Ethics“ kann sich also das Problem ergeben, dass Ethik in KI-Anwendungen nicht effektiv umgesetzt wird.

Über Ansätze, das zu überwinden, sollte nachgedacht werden. Ich sehe dabei folgende Möglichkeiten:

- Die Sprachfähigkeit in Sachen KI-Ethik soll für alle Beteiligte verbessert werden. Dazu gehört, sich mit Grundlagen der Ethik im Kontext von Informatik und KI zu beschäftigen. Hier gibt es zwar mittlerweile reichlich Literatur [u.a. Thomas Matzner, Informatikethik], jedoch wurde das Wissen vieler Akteure noch in einer Zeit geprägt, zu dem so etwas nicht einmal zum optionalen Ausbildungs-Curriculum gehörte.
- Darauf aufbauend sollte dann jeder Akteur seinen eigenen ethischen Standpunkt reflektieren und auch erkennen können, welche ethischen Ansätze seine Dialogpartner in einer Debatte oder einem Projekt verfolgen. Dadurch ist im Idealfall gegenseitige Ergänzung möglich.
- Bei alledem sollte auch immer im Blick behalten werden, dass es ethische „Blind Spots“ gibt, die auch bei technisch einwandfreien KI-Systemen auftreten können, selbst wenn sie - um alle Kritiker zu befriedigen - zugleich ressourcenschonend und diskriminierungsfrei wären: Beispielsweise wie es ethisch zu beurteilen ist, wenn Menschen aufgrund im Alltag etablierter KI-Anwendungen Fähigkeiten „verlernen“ und so übermäßig abhängig von KI werden.

-----  
Angaben zum Autor:

Tobias Hamberger, Diplom-Informatiker (TU München), ist seit über zwei Jahrzehnten in einem mittelständischen Unternehmen (MicroNova AG) tätig mit dem Schwerpunkt Embedded Testing und Functional Safety. Daneben ist dort auch KI ein Consulting-Schwerpunkt.

Ehrenamtlich arbeitet er in der „Taskforce on Artificial Intelligence and Intellectual Property“ der Episcopal Church (USA) mit. Aus der dortigen Beschäftigung mit ethischen Fragen entstand ein Teil der hier geäußerten Gedanken.

*[Im Folgenden wird der Text noch einmal auf Englisch aufgeführt.]*

--[ENGLISH:]--

Why we should abandon the term „Artificial Intelligence“?

The opening article already provided an answer to this question. Among other things, it was criticized that the term "Artificial Intelligence" creates misconceptions and false expectations about this topic (with technical and ethical implications).

In my article, I specifically examine the term "AI ethics," which could be somewhat vague, as different actors in the field of "AI" could have different understandings. Therefore, this term should also be viewed critically.

Because, if participants have completely different understandings of AI ethics without being aware of it, this can lead to disappointment and mutual alienation, when instead, collaboration and mutual enrichment of ethical perspectives, would be needed.

In addition to pertinent examples where the term "AI Ethics" is misused for "Whitewashing" purposes, examples include:

- Due to ecological and societal consequences of the use of AI, many expect AI Ethics to effectively mitigate these consequences. While ethical considerations may then lead to principles such as "fairness" or "freedom from bias," which an AI application must fulfill, on the one hand, implementation often becomes lost in technical details, and the underlying social problem is lost sight of. On the other hand, the (possibly excessive) expectations of critics remain unfulfilled, reinforcing their rejection of "AI."
- It is commendable that many large companies have already established policies and principles for dealing with AI, which, on paper, appear very solid. On the one hand, this leads to increased public trust. On the other hand, however, it can also be observed that policies created by companies themselves are non-binding. Therefore, this trust can be lost again through arbitrary alterations.

Different interpretations of the term "AI Ethics" can therefore lead to the problem that in the end, ethics is not effectively implemented in AI applications. Approaches to overcome this should be considered. I see following possibilities:

- One goal is to improve the literacy of all participants regarding AI ethics. This includes engaging with the fundamentals of ethics in the context of computer science and AI. While there is nowadays a rich choice of literature on this topic, the knowledge of many participants was formed at a time when such topics were not even optional part of the education curriculum.
- Based on this, each participant should then be able to reflect their own ethical standpoint and likewise recognize which ethical approaches their dialogue partners are pursuing in a debate or project. In the ideal case, different approaches should complement each other.
- It should also be kept in mind that there are ethical "blind spots" that can occur even in technically flawless AI systems, even if they were – to satisfy all critics – simultaneously sustainable, fair, and free from bias: For example, how to judge ethically if people "unlearn" skills due to AI applications established in everyday life and thus become overly dependent on AI.