

Nachhaltigkeit mitdenken im gesamten Lebenszyklus von KI

Erarbeitet von

Marie von Lobenstein, M.A.

Das Thema „Nachhaltige KI“ ist äußerst komplex und wird aus verschiedenen Fachdisziplinen und Perspektiven betrachtet. In diesem Video haben wir daher eine Auswahl getroffen, um einen ersten Überblick zu geben. Eine vollständige Behandlung des Themas würde den Rahmen der Lehrveranstaltung sprengen und ist im Rahmen dieses Kurses nicht möglich.

Lernziele	1
Inhalt	2
Einstieg	2
Der Lebenszyklus eines KI-Systems	2
Der Lebenszyklus von KI und Nachhaltigkeit	3
Beispiel 1: Der rassistische Seifenspende	4
Beispiel 2: Das ultimative Modell und Ressourcenverbrauch	4
Take-Home Message	5
Quellen	6
Weiterführendes Material: Lebenszyklus KI	7
Disclaimer	7

Lernziele

- Du kannst erklären was gemeint ist, wenn die Rede vom Lebenszyklus eine KI ist
- Du kannst erklären, warum wir den gesamten Lebenszyklus von KI betrachten müssen, um die Nachhaltigkeit von KI zu bewerten
- Du kannst Beispiele für Nachhaltigkeitsbetrachtungen in Verbindung mit einem KI-Lebenszyklus nennen

Inhalt

Einstieg

Herzlich willkommen zu unserem Video über die Nachhaltigkeit von KI-Systemen, ein Blick auf den Lebenszyklus einer KI.

Wir haben uns im Video zum Thema Auswirkungen (Link) ja bereits angesehen, dass die Frage „wie wirkt sich eine KI auf unsere Gesellschaft und unsere Umwelt aus?“ wichtig ist, wenn wir bewerten wollen, wie nachhaltig ein KI-System ist.

Im white paper der EU-Kommission aus 2020 wird ein weiterer Aspekt deutlich, der in unseren Nachhaltigkeitsbetrachtungen zur Auswirkungsebene hinzukommt:

„[...] Angesichts der zunehmenden Bedeutung der KI müssen die ökologischen Auswirkungen von KI-Systemen auf die Umwelt während ihres **gesamten Lebenszyklus** und über die gesamte Lieferkette hinweg gebührend berücksichtigt werden, z. B. im Hinblick auf den Ressourcenverbrauch für das Training von Algorithmen und die Speicherung von Daten.“

Quelle [1] Eigene Übersetzung

In diesem Video werden wir also einen genaueren Blick auf den Lebenszyklus eines KI-Systems werfen. Hier ist die zentrale Frage: Warum müssen wir uns den gesamten Lebenszyklus eines KI-Systems ansehen, wenn wir über Nachhaltigkeit sprechen?

Der Lebenszyklus eines KI-Systems

Der Lebenszyklus von Künstlicher Intelligenz umfasst verschiedene Phasen, die von der Konzeption und Entwicklung bis hin zur Nutzung und möglicherweise zur Entsorgung oder Wiederverwendung des Systems reichen. Dieser Lebenszyklus wird in der Wissenschaft mal mehr und mal weniger detailliert aufgeschlüsselt.

Quelle [2][3][4][5]

Für den Zweck dieses Videos werde ich euch eine vereinfachte Version dieses Lebenszyklus zeigen. Wenn ihr euch detaillierter dazu informieren wollt, dann schaut euch die Leseempfehlung am Ende des Videos an.

Hier wird eine Lebenszyklus-Grafik eingebaut.

Zu Beginn des Lebenszyklus einer KI steht die Konzeptionierung, also was wird gebraucht, wie setzen wir das um, etc. In dieser Phase der „[...] Model Requirements“ entscheiden die Designer*innen beispielsweise, welche Funktionen mit maschinellem Lernen implementiert werden können [...]. Am wichtigsten ist, dass sie in dieser Phase entscheiden, welche Arten von Modellen für das jeweilige Problem am besten geeignet sind."

Quelle [6] Eigene Übersetzung

In den Datenphasen des Zyklus werden verfügbare Datensätze integriert oder Daten gesammelt. Anschließend werden die Daten gesäubert und gelabelt, wobei beispielsweise Ungenauigkeiten bereinigt und Daten beschriftet werden.

Darauf folgen mehrere Modellphasen, in denen das gewählte Modell mit den gereinigten und beschrifteten Daten trainiert, getestet und evaluiert wird.

Zum Schluss wird das Modell dann eingesetzt und kontrolliert. Der Einfachheit halber ist das Modell hier linear dargestellt. Ihr solltet jedoch wissen, dass es in einem KI-Lebenszyklus mehrere Feedback-Loops gibt und der Machine-Learning-Workflow meist nicht-linear verläuft. Phasen, wie die Modell Evaluation oder die Kontrolle des Systems beispielsweise, passieren im Optimalfall zeitgleich und mehrfach in verschiedenen Zyklusphasen.

Quelle [7]

Wir nutzen KI-Anwendungen einfach mal so nebenbei im Alltag. Diese Anwendungen sind allerdings aus einem langen und aufwendigen Prozess hervorgegangen, in dessen verschiedenen Phasen wir unter anderem die Nachhaltigkeit berücksichtigen müssen. Wir schauen uns also zusätzlich dazu, wie sich das KI-System auswirkt, auch die einzelnen Phasen im Entstehungsprozess dieser KI an, um die Nachhaltigkeit des KI-Systems zu bewerten.

Der Lebenszyklus von KI und Nachhaltigkeit

Damit wir bewerten können, ob ein KI-System nachhaltig ist, müssen wir also den ganzen Lebenszyklus betrachten. „Denn nur so können KI-Systeme entworfen, entwickelt und eingesetzt werden, die nicht auf der einen Seite Probleme lösen und auf der anderen Seite wieder neue Probleme schaffen“,

Quelle [8]

sich also zum Beispiel indirekt auf ein SDG-Ziel auswirken. Beispielsweise kann eine unzureichende Datenerfassung zu Ungleichheiten führen, während ein übermäßiger Ressourcenverbrauch bei der Modellentwicklung die Umwelt belasten kann. Daher ist es von größter Bedeutung, den Lebenszyklus der KI kritisch zu hinterfragen und nachhaltige Entscheidungen zu treffen, um eine positive Wirkung auf die Gesellschaft und Umwelt zu erzielen.

Ich veranschauliche das hier für euch mal mit zwei Beispielen:

Beispiel 1: Der rassistische Seifenspender

2019 erschien ein Video im Internet, auf dem folgendes zu sehen ist: Ein Schwarzer Mann versucht auf der Toilette in einem Hotel den automatisierten Seifenspender zu benutzen. Dieser erkennt aber seine Hand unter dem Sensor nicht. Als sein Freund, ein weißer Mann, es ihm nachmacht, kommt die Seife heraus. Die Kommentare unter dem Video: Ist dieser Seifenspender rassistisch? Die Antwort: nein. Der Seifenspender ist nicht rassistisch, der Seifenspender führt die Funktion aus, für die er programmiert wurde, das aber eben nicht besonders effektiv für alle. Was ist also passiert?

Der Seifenspender funktioniert durch Nah-Infrarot-Technologie. Man hält die Hand unter den Sensor, und das ausgeworfene Licht wird von der Haut zurückgeworfen und löst den Sensor aus. Dunklere Hauttöne absorbieren mehr Licht, sodass nicht genug Licht zum Sensor zurück reflektiert wird, um den Seifenspender zu aktivieren. Das bedeutet, dass beispielsweise Schwarze Toilettenbenutzer*innen diesen Sensor nicht auslösen.

Quelle [9]

Dieses Beispiel soll euch verdeutlichen, dass Technologien mit der besten Intention programmiert werden können, also ein Problem zu lösen, aber auch gleichzeitig Probleme schaffen, wenn im gesamten Lebenszyklus der Anwendung nicht auf Aspekte wie eben Nachhaltigkeit geachtet wird. Wo im Lebenszyklus dieser Anwendung ist, also was schiefgelaufen oder hätte eingegriffen werden können?

Der Seifenspender wurde von einem Unternehmen namens Technical Concepts entwickelt, das unbeabsichtigt einen diskriminierenden Seifenspender herstellte. Wie kann es zu dieser Fehlentwicklung kommen? Zum Beispiel, weil niemand im Unternehmen daran dachte, sein Produkt an Menschen mit dunkler Hautfarbe zu testen. Es kann auch sein, dass der Algorithmus an gebiaseten Daten trainiert wurde. Möglich ist auch, dass das Entwicklungsteam nicht divers genug gewesen ist und somit niemand auf die Idee gekommen ist, den Sensor auch auf Schwarzer Haut zu testen. Vielleicht wurde auch nicht genau genug hingesehen in der Phase der Model Evaluierung.

Dies sind nur mögliche Erklärungen für die Fehlfunktion des Seifenspenders. Was genau schiefgelaufen ist, wissen wir nicht.

Ich habe euch noch ein zweites Beispiel mitgebracht.

Beispiel 2: Das ultimative Modell und Ressourcenverbrauch

Wenn man sich zu Nachhaltigkeit von KI informiert, stößt man schnell auf Debatten rund um den Energieverbrauch von KI.

Um diese Problematik zu veranschaulichen, werden dann meist CO2 Berechnungen genutzt, die zeigen sollen, wie viel Energie das Training eines KI-Modells verbraucht. Wie z. B.:

„Für das Training eines großen Sprachmodells wie GPT-3 werden schätzungsweise knapp 1.300 Megawattstunden (MWh) Strom verbraucht; das ist etwa so viel, wie 130 US-Haushalte jährlich verbrauchen. Zum Vergleich: Für das Streaming einer Stunde Netflix werden etwa 0,8 kWh (0,0008 MWh) Strom benötigt. Das bedeutet, dass man 1.625.000 Stunden Netflix schauen müsste, um die gleiche Menge an Strom zu verbrauchen, die für die Ausbildung des GPT-3 benötigt wird.“

Quelle [10]

Bei diesen Berechnungen muss man allerdings unter anderem folgendes beachten:

Komplexität der Berechnung: Ja, ein KI-Modell zu trainieren, kann wirklich viel Energie verbrauchen. Diese Verbrauchsberechnungen können jedoch auch zu kurz greifen, wenn sie als Anhaltspunkt dafür dienen sollen, ob eine KI-Anwendung nachhaltig ist oder nicht. Denn, sie betrachten nicht den ganzen Lebenszyklus einer KI. Wenn wir uns den ganzen Lebenszyklus ansehen, bedeutet das, dass sowohl der laufende Energieverbrauch als auch die Umweltauswirkungen der KI während ihres gesamten Lebenszyklus – von der Datenerfassung bis zum Betrieb des endgültigen Modells – berücksichtigt werden müssen. Die Häufigkeit des Trainings sowie der Umfang der einzelnen Schritte bei der Entwicklung der KI sind wichtige Faktoren. Wir müssen auch über die gesamte Lebensdauer der KI-Software und -Hardware nachdenken, einschließlich der Art und Weise, wie sie hergestellt und verwendet werden. Es stellen sich wichtige Fragen wie „Welche Materialien werden für die Hardware verwendet?“, „Wie viele Trainings eines KI-Systems sind in einem Ressourcen Kosten- und Nutzen Vergleich vertretbar?“, „Ist ein KI-System notwendig, um ein Problem zu lösen?“, oder „Gibt es vielleicht ressourcenschonendere Ansätze?“.

Vergleichbarkeit: Es ist wichtig zu beachten, dass verschiedene KI-Modelle unterschiedliche Trainingsanforderungen haben. Ein Modell könnte weniger Energie benötigen, um trainiert zu werden, aber möglicherweise auch weniger leistungsfähig sein. Daher ist ein direkter Vergleich des Energieverbrauchs möglicherweise nicht aussagekräftig, wenn nicht auch die Leistung berücksichtigt wird.

Optimierung und Effizienz: Die Effizienz von KI-Modellen variiert stark je nach Architektur, Algorithmen und Optimierungstechniken. Ein schlecht optimiertes Modell kann erheblich mehr Energie verbrauchen als ein gut optimiertes Modell mit vergleichbarer Leistung. Es ist daher wichtig, nicht nur den absoluten Energieverbrauch zu betrachten, sondern auch die Effizienz des Modells in Bezug auf die erzielte Leistung zu bewerten.

Take-Home Message

Wenn wir die Nachhaltigkeit eines KI-Systems betrachten, sollten sich diese Betrachtungen nicht auf Teilaspekte eines KI-Systems beschränken, sondern sollten den gesamten Lebenszyklus dieses Systems unter die Lupe nehmen. Nur so können wir sicherstellen, dass KI-Systeme nicht nur Probleme lösen, sondern auch keine neuen Probleme schaffen. Die Nachhaltigkeit von KI-Systemen erfordert also eine ganzheitliche Betrachtung ihres Lebenszyklus. Im Bereich Gesellschaft kann das unter anderem bedeuten, auf ein diverses Entwicklungsteam zu achten, auf korrekte Datensätze und korrekte Evaluierungen. Im

Bereich Ressourcenverbrauch genügt es nicht, lediglich den Energieverbrauch während des Trainings eines KI-Modells zu analysieren. Wir müssen auch die Phasen der Datenerfassung, Modellentwicklung, Implementierung und Entsorgung berücksichtigen. Dieser umfassende Ansatz ermöglicht es uns, die Auswirkungen der KI auf die Umwelt, die Gesellschaft und die Wirtschaft besser zu verstehen und sicherzustellen, dass sie langfristig nachhaltig ist. Durch die Einbeziehung verschiedener Aspekte wie Ressourcenverbrauch, Effizienz und soziale Auswirkungen können wir fundierte Entscheidungen treffen, um eine positive Wirkung auf unsere Welt zu erzielen.

Was wir uns in diesem Video noch nicht angeschaut haben, ist, wer eigentlich Einfluss darauf hat, die Nachhaltigkeit eines KI-Systems mitzugestalten und zu entscheiden. Wer oder was kann eigentlich Entscheidungen treffen, im und über den KI-Lebenszyklus? Wer oder was kann beeinflussen, wie sich ein KI-System auswirkt? Außerdem stellt sich die Frage, auf welchen Ebenen finden diese Entscheidungen und Gestaltungsmöglichkeiten statt? Was es mit diesen Ebenen und Einflussnehmenden auf sich hat, seht ihr im Video nachhaltige KI und Wirkungsebenen.

Quellen

- Quelle [1] E Commission. (2020). On artificial intelligence—a European approach to excellence and trust.
- Quelle [2] Shearer C (2000) The CRISP-DM model: the new blueprint for data mining. *J Data Warehous* 5(4):13–22
- Quelle [3][6]
- S. Amershi *et al.*, "Software Engineering for Machine Learning: A Case Study," *2019 IEEE/ACM 41st International Conference on Software Engineering: Software Engineering in Practice (ICSE-SEIP)*, Montreal, QC, Canada, 2019, pp. 291-300, doi: 10.1109/ICSE-SEIP.2019.00042.
- Quelle [4] van Wynsberghe, A. Sustainable AI: AI for sustainability and the sustainability of AI. *AI Ethics* 1, 213–218 (2021). <https://doi.org/10.1007/s43681-021-00043-6>
- Quelle [5] De Silva, D., & Alahakoon, D. (2022). An artificial intelligence life cycle: From conception to production. *Patterns*, 3(6).
- Quelle [7] Haakman, M., Cruz, L., Huijgens, H. *et al.* AI lifecycle models need to be revised. *Empir Software Eng* 26, 95 (2021). <https://doi.org/10.1007/s10664-021-09993-1>
- Quelle [8] Rohde, F., Wagner, J., Reinhard, P., Petschow, U., Meyer, A., Voß, M., & Mollen, A. (2021). Nachhaltigkeitskriterien für künstliche Intelligenz. *Schriftenreihe des IÖW*, 220, 21.

Quelle [9][10] Synclair Goethe, T. (2019). *Bigotry encoded: Racial bias in technology*. Reporter.
<https://reporter.rit.edu/tech/bigotry-encoded-racial-bias-technology>

Weiterführendes Material: Lebenszyklus KI

Silva, D., & Alahakoon, D. (2022). An artificial intelligence life cycle: From conception to production. *Patterns*, 3(6).

Haakman, M., Cruz, L., Huijgens, H. *et al.* AI lifecycle models need to be revised. *Empir Software Eng* **26**, 95 (2021). <https://doi.org/10.1007/s10664-021-09993-1>

Shearer C (2000) The CRISP-DM model: the new blueprint for data mining. *J Data Warehous* 5(4):13–22

Van Wynsberghe, A. Sustainable AI: AI *for* sustainability and the sustainability *of* AI. *AI Ethics* **1**, 213–218 (2021). <https://doi.org/10.1007/s43681-021-00043-6>

Disclaimer

Transkript zu dem Video „KI und Nachhaltigkeit: Nachhaltige KI- Der Lebenszyklus“, Marie von Lobenstein. Dieses Transkript wurde im Rahmen des Projekts ai4all des Heine Center for Artificial Intelligence and Data Science (HeiCAD) an der Heinrich-Heine-Universität Düsseldorf unter der Creative Commons Lizenz [CC-BY 4.0](https://creativecommons.org/licenses/by/4.0/) veröffentlicht. Ausgenommen von der Lizenz sind die verwendeten Logos, alle in den Quellen ausgewiesenen Fremdmaterialien sowie alle als Quellen gekennzeichneten Elemente.