BIGDATA INSIDER

Künstliche Intelligenz

KI muss vertrauenswürdig, fair, sicher und erklärbar sein

23.08.2021 | Autor / Redakteur: Michael Matzer / Nico Litzel

Eine Studie von Morning Consult hat untersucht, welche Bedingungen für die Einführung von KI-Technologien und -Lösungen im Business erfüllt sein müssen. Einer der wichtigsten – mit 91 Prozent – ist die Vertrauenswürdigkeit von KI-Ergebnissen. Andrea Martin, die Leiterin des IBM Watson Centers Munich, liefert dazu klare Antworten.

Andrea Martin, Leiterin des IBM Watson Center Munich und Mitglied des KI-Rates der Bayerischen Staatsregierung (Bild: IBM)

Andrea Martin, Leiterin des IBM Watson <https://www.bigdata-insider.de/was-ist-watson-a-572251/> Center Munich und ehemalige CTO für die DACH-Region, war zwischenzeitlich Mitglied der KI-Enquete-Kommission des Deutschen Bundestags und ist inzwischen Mitglied des KI-Rates der Bayerischen Staatsregierung. „Das Vertrauen unserer B2B-Kunden in unsere KI-Technologie Watson ist uns sehr wichtig, nicht nur in die KI, sondern auch allgemein in die Technologie." Sie zitiert Ginni Rometty, die vormalige IBM-CEO, die 2017 in Davos gesagt hatte: „Wenn Sie leistungsstarke Technologien in diese Welt einführen, haben Sie die Verantwortung, dass dies auf die richtige Weise geschieht."

Dieser Wunsch nach Vertrauen werde durch aktuelle Studien bestätigt. So etwa durch die Global AI Adoption Studie mit Morning Consult <https://filecache.mediaroom.com/mr5mr_ibmnewsroom/191468/IBM's Global AI Adoption Index 2021_Executive-Summary.pdf> . Diese Studie mit 5.500 befragten Unternehmen in aller Welt hat einerseits untersucht, wo Unternehmen KI einsetzen oder den Einsatz planen und mit welchen Zielen und andererseits auch, wo hier die größten Herausforderungen liegen.

„In Deutschland", so Martin weiter, „sagen 74 Prozent der IT-Profis und 88 Prozent der Unternehmen, in denen KI bereits eingesetzt wird, dass Vertrauen wichtig oder sehr wichtig ist. Also, dass man sich darauf verlassen kann, dass die Ergebnisse der KI fair, zuverlässig und sicher sind. Für 85 Prozent ist es wichtig, dass man erklären kann, wie eine

KI zu ihrem Ergebnis gekommen ist." Wenn man die deutschen Zahlen mit dem weltweiten Niveau vergleicht, liegt Deutschland ziemlich im Durchschnitt.

BILDERGALERIE

„Fair, zuverlässig und sicher" seien Themen, an denen IBM schon lange arbeite und die auch in der Enquete-Kommission immer wieder diskutiert worden seien. Besucher könnten das auch im IBM Watson Center Munich sehen. „Wir haben eine Demo zum Thema Fairness bzw. Unvoreingenommenheit in KI", berichtet Martin und fährt fort: „Wir arbeiten z. B. mit dem Bayerischen Forschungsinstitut, mit Fortiss <https://www.fortiss.org/forschung/forschungsfelder/detail/center-for-ai> am digitalen Testfeld Autobahn zusammen. So etwa, wenn es um die Erklärbarkeit von Spurwechseln beim autonomen Fahren anhand von digitalen Zwillingen geht. Die KI verwertet Informationen, die die Insassen mitunter noch nicht sehen oder wahrnehmen können. Dann ist es wichtig zu verstehen, warum das Auto von der linken auf die rechte Spur wechselt, weil es weiß, dass hinter der nächsten Kuppe ein Stau beginnt und es langsamer fahren muss. Darum geht es in dieser Situation." Die Erklärbarkeit dieser Verhaltensweise ist ebenso wichtig wie die Transparenz. Das sei der Anwendungsfall, den IBM zeige. „Der Zweck von KI besteht dazu, die Fähigkeiten des Menschen zu erweitern, also Augmented Intelligence <https://www.bigdata-insider.de/was-ist-augmented-intelligence-a-733360/> ", resümiert Martin.

Der AI Adoption Report nennt Sprachverarbeitung (NLP <https://www.bigdata-insider.de/was-ist-natural-language-processing-a-590102/>) als beliebteste Anwendung von KI. „Fast die Hälfte der befragten Unternehmen nutzt bereits NLP-Anwendungen, wie etwa Chatbots <https://www.bigdata-insider.de/was-ist-ein-chatbot-a-690591/> . Und eines von vier Unternehmen will NLP in den nächsten zwölf Monaten einführen. Kundendienst ist der wichtigste Use Case, denn er wurde von 52 Prozent der befragten IT-Mitarbeiter genannt: Chatbots und Co. sollen den Kundenservice verbessern, insbesondere vor dem Hintergrund von Covid-19.

Zu den größten Hindernissen bei der KI-Nutzung zählten die Befragten den Zugang zu den benötigten Daten. Mehr als zwei Drittel der IT-Mitarbeiter zapfen Daten aus mehr als 20 Quellen an. Nahezu 90 Prozent der Befragten gaben daher an, dass die Fähigkeit, ihre KI-Projekte ausführen zu können, egal wo die Daten lägen, der Schlüssel für die Einführung der KI-Technologie sei. Weitere Hürden sind bereits bekannt: der Mangel an einschlägigen Fachkenntnissen (39 Prozent), Datensilos und zunehmende Datenkomplexität (32 Prozent) sowie der Mangel an Entwicklungsplattformen und -werkzeugen für KI-Modelle (28 Prozent).

KI Toolkits

Auf der technischen Ebene stehen nach Angaben von Andrea Martin für Machine-Learning <https://www.bigdata-insider.de/was-ist-machine-learning-a-592092/> -Modell-Toolkits zur Verfügung, die helfen, Empfehlungen von KI-Lösungen zu erklären. Als erstes schaffen „AI Factsheets" Transparenz über den Zweck des KI-Lösung, verwendete Daten, die Modelle etc. und schaffen so eine erste Grundlage für die Akzeptanz der Lösung. Zweitens sind „AI Explainabilty 360" und „AI Fairness 360" Open Source Toolkits, deren Entwicklung von IBM Research initiiert wurde, mit zahlreichen Algorithmen, die ML- Modelle interpretieren und erklären bzw. hinsichtlich Fairness überprüfen. „Diese Open-Source-Toolkits kommen zwar von IBM Research, aber sie werden durch die Open Source Community <https://www.ibm.com/opensource/> weiterentwickelt, die wir dazu einladen." Lokale User Groups würde ihre eigenen Präferenzen und Prioritäten einbringen.

Und drittens bietet IBM noch Watson OpenScale <https://www.ibm.com/de-de/cloud/watson-openscale> an. „Das ist ein Service, den man aus der Cloud beziehen, aber auch im eigenen Rechenzentrum verwenden kann – mit IBM und Nicht-IBM-KI-Lösungen", so Martin. „Der Service unterstützt Nutzer dabei, KI-Lösungen schon während des Betriebs zu analysieren und zu überwachen."

Damit könnten frühzeitig Probleme erkannt werden, aber auch im laufenden Betrieb Ergebnisse erklärt und auf ihre Fairness überprüft werden. Zum Beispiel könne man in Dashboards <https://www.bigdata-insider.de/was-ist-ein-business-intelligence-dashboard-a-581644/> sichtbar machen, welche Eingabeparameter Ergebnisse wie stark beeinflussen. „Die Analyse sagt auch aus, wenn es zu wenige Daten gibt, um eine Aussage zuverlässig zu machen. Dann kann der Nutzer entscheiden, auf die KI-Aussage zu verzichten und vielmehr den entsprechenden Fall quasi manuell zu betrachten." „Ich kann sagen: KI ist eine Entwicklung, die auf weltweiten Ressourcen beruht."

IBM-Grundprinzipien

Das IBM Watson Center Munich spielt dabei eine koordinierende Rolle. „Der Schwerpunkt liegt u. a. auf der Kooperation mit Fortiss", erläutert Martin. „Unsere Projektvorgehensweisen sind immer eine gemeinschaftliche Arbeit von Kunden- und IBM-Expertinnen und -Experten. Das gilt natürlich auch im KI-Umfeld. Methoden wie Design Thinking oder unsere Garage-Methodik <https://www.ibm.com/de-de/garage> sehen vor, dass wir gemeinsam das Ziel des Projekts definieren, die gewünschten Ergebnisse festlegen, die zu unterstützenden Geschäftsprozesse betrachten sowie die vorhandenen und benötigten Daten, etc. Weiterhin beinhalten die von uns verwendeten Methoden, dass bereits während der Design- und Entwicklungsphase auf das Thema Vertrauenswürdigkeit geachtet wird."

Sicherheit und Datenschutz haben Priorität, und die Einhaltung von Gesetzen sei Pflicht, erklärt Martin. „Es sind wesentliche Elemente einer KI-Lösung, dass wir Datenschutz und Datensicherheit einhalten." Zum anderen sei das IBM-Geschäftsmodell nicht auf Daten aufgebaut. „Das heißt, wir sind ein Technologieunternehmen und alle Daten, die für eine KI-Lösung benötigt werden und alle Erkenntnisse, die mithilfe von KI gewonnen werden, gehören

demjenigen, der die Daten erzeugt hat. Das ist eines unserer wichtigsten Prinzipien <https://www.ibm.com/blogs/policy/trust-principles/> ."

Aktuell arbeitet IBM mit Partnern an der Initiative „responsible.computing()", die den verantwortungsvollen Umgang mit Informationstechnologie von einer systemischen Seite anhand von sechs miteinander verbundenen Bereichen betrachtet. Darüber hinaus werden auch bereichsübergreifende Themen wie Nachhaltigkeit, Klima, Vielfalt, Ethik, genereller und barrierefreier Zugang, Offenheit gegenüber neuen Technologien sowie deren Sicherheitsaspekte beleuchtet. Unternehmen können mithilfe eines Tests, dem sogenannten Maturity Assessment, ihren Reifegrad selbst einschätzen, inwieweit sie verantwortungsvoll mit neuen Technologien umgehen. Aus den Testergebnissen können Unternehmen dann ableiten, wo sie stehen, und entscheiden, was ihre nächsten konkreten Schritte sind.

IBMs Ethik-Rat

„Es gibt auch Projekte, die wir nicht machen", berichtet Martin. „Für solche Entscheidungen haben wir ein KI-Ethik-Board (AI Ethics Board) <https://www.ibm.com/de-de/artificial-intelligence/ethics> , das global agiert." Das IBM AI Ethics Board sei eine Weiterentwicklung des ethischen Entscheidungsfindungsprozesses, den IBM auf KI anwendet. Das Board wurde als zentrales, interdisziplinäres Gremium gegründet, um eine Kultur ethischer, verantwortungsvoller und vertrauenswürdiger KI bei IBM zu unterstützen.

„Unsere Zielsetzung ist die Unterstützung eines zentralen Governance-, Prüfungs- und Entscheidungsfindungsprozesses für IBMs Richtlinien, Verfahren, Kommunikation, Forschung, Produkte und Services im Bereich Ethik. Durch Aufnahme unserer langjährigen Grundsätze und ethischen Denkweisen ist das Gremium ein Mechanismus, mit dem IBM unser Unternehmen und alle IBM-Mitarbeiterinnen und -Mitarbeiter für unsere Werte verantwortlich macht."

Rekrutierung

Die Morning-Consult-Studie besagt, dass begrenzte KI-Expertise oder -Wissen mit 33 Prozent eines der größten Hindernisse für die Einführung von KI in Europa sei, gefolgt von zunehmender Datenkomplexität und Datensilos mit 29 Prozent. In Deutschland sind die Zahlen sogar noch höher: Die begrenzte KI-Expertise wird sogar zu 37 Prozent genannt, die Datenkomplexität und Datensilos sogar zu 38 Prozent. „Damit sind wir negativer Spitzenreiter in Europa", kommentiert Martin und erläutert, was IBM auf dem Feld der KI dagegen unternimmt.

„Deshalb haben Schulungen und Personalentwicklung für deutsche IT-Experten bei der Einführung von KI Priorität", so Martin. Immerhin: 39 Prozent der Befragten haben laut Studie angegeben, im nächsten Jahr in diesen Bereich investieren zu wollen. „Aber natürlich können hier auch Technologie-Unternehmen beratend und durch Projektunterstützung zur Seite stehen. Wir haben dafür viele KI-Expertinnen und -Experten in München und anderswo."

(ID:47541490)