



Intercultural school
Talents pour le monde

EPREUVES D'ADMISSION

4 CIT

SESSION DE JANVIER 2020

LANGUE DE TRAVAIL : ALLEMAND

Épreuve d'admission en 4^{ème} année CIT

Traduction ALL>FR

Aucun document autorisé, traduire les passages entre gras et en crochets

[Künstliche Intelligenz und Ethik

Interview: Eine Frage der Ethik

Künstliche Intelligenz verbessert unseren Alltag erheblich. Aber damit der Mensch der smarten Technologie vertrauen kann, muss er ihr zuallererst ethische Entscheidungsgewalt zutrauen. In München erforschen Wissenschaftler deshalb an einem neu gegründeten Institut die Zusammenhänge zwischen Ethik und Künstlicher Intelligenz. Ein Gespräch mit Christoph Lütge, Leiter des Institute for Ethics in Artificial Intelligence (IEAI).

Herr Professor Lütge, kann man Maschinen beibringen, was ein ethisches Verhalten ist?

Die Maschine handelt aufgrund eines Programms. Aber in dieses Programm fließen natürlich bestimmte Einstellungen des Programmierers ein, außerdem die Daten, mit denen das System gefüttert wird. Dann muss man sich klar machen, dass wir es bei der Künstlichen Intelligenz mit einem Ablauf zu tun haben, der nicht mehr deterministisch ist. Die Machine Learning Prozesse führen vielmehr dazu, dass die Resultate nicht immer im Detail voraussehbar und leicht nachvollziehbar sind. Eine wichtige Aufgabe wird sein, diese Prozesse zumindest im Prinzip erklärbar zu machen: Warum hat das System so und so entschieden? Wie kam es dazu? Das ist eine große Herausforderung.

Wie wollen Sie mit Ihrem Institut dieser Herausforderung begegnen?

Wir loten Künstliche Intelligenz in all ihren ethischen Facetten aus und konzentrieren uns derzeit auf acht verschiedene Forschungsprojekte. Eines beschäftigt sich beispielsweise mit der Ethik des autonomen Fahrens. Ein anderes dreht sich um KIbasierte Entscheidungshilfen im Klinik-Alltag. Alle Forschungsprojekte werden von je einem Informatiker und einem Ethiker oder einer Sozialwissenschaftlerin geleitet. Auf diese Weise stellen wir sicher, dass die Perspektiven aus verschiedenen Disziplinen in die Forschung einfließen. Es geht darum, ethische Probleme bei Künstlicher Intelligenz zu benennen, bewusst zu machen und in die Gesellschaft zu tragen. Aber wir wollen es nicht bei einem klassischen Forschungoutput in Form von Papers belassen, sondern auch Richtlinien, Standards und Normen für künftige KI-Systeme entwickeln.

Eine Art TÜV¹ für Künstliche Intelligenz also?

Da wir keine Zertifizierungsagentur sind, würde ich den Begriff TÜV wohl nicht verwenden. Aber es kann durchaus in die Richtung gehen, dass wir Entscheidungsträgern aus Politik, Wirtschaft und Zivilgesellschaft Vorschläge unterbreiten, die klarstellen, an welchen ethischen Anhaltspunkten sich ein geplantes KI-System orientieren sollte.]

Gibt es denn faire und vorurteilsfreie Algorithmen?

Die einfache Antwort wäre zu sagen: Ja. Aber ich will etwas vorsichtiger sein. Es kann natürlich passieren, dass in solchen Systemen bestimmte Gruppen systematisch benachteiligt werden. Eine Firma entwickelte etwa eine Software, die Bewerbern automatisch passende Jobangebote unterbreiten sollte. Der Algorithmus hatte aber dafür gesorgt, dass Frauen schlechtere Jobempfehlungen erhielten. Genauso ist bekannt, dass manche Algorithmen, die die Rückfallgefahr von Straftätern in den USA berechnen, Afroamerikaner diskriminieren.

Wie kann man Künstliche Intelligenz ethisch auf Kurs bringen?

Man darf es sich nicht so vorstellen, dass man "unfaire" Elemente aus dem System entfernt, und dann ist alles gut. Um einen Algorithmus zu ändern, muss man innerhalb des KI-Systems möglichst viele Szenarien durchspielen und beobachten, wann die Änderungen an den Randbedingungen zu einem fundamental anderen Resultat führen. Auf dieser Grundlage kann man den Algorithmus nachjustieren. Das ist ein Prozess, der wirklich alles andere als trivial ist, und in erster Linie gilt es zu vermeiden, dass so etwas überhaupt nötig ist. Zahlreiche KI-Systeme helfen aber bereits heute, faire Entscheidungen zu treffen.

Zum Beispiel?

Hier ist eine Studie zu Richtern sehr anschaulich. Sie wurde von israelischen und US-amerikanischen Forschern vor einigen Jahren durchgeführt und kommt zu dem Ergebnis, dass Richterinnen und Richter vor und nach dem Essen anders entscheiden. Wer also das Pech hat, sein Urteil vor dem Mittagstisch zu erhalten, erhält unter Umständen eine härtere Strafe als von einem sattem Richter. In KI-Systemen spielt die Auswirkung von Hunger keine Rolle. Auch autonome Fahrzeuge werden künftig zahlreiche Unfälle verhindern. Dennoch wird es zu Kollisionen kommen, die auf Fehlfunktionen der KI zurückzuführen sind. Die Frage ist, wie man mit diesen Situationen umgeht. Deswegen müssen wir ethische Regeln finden, um die KI-Systeme noch weiter zu verbessern. Und wir müssen klären, wie Unternehmen reagieren, wenn es aufgrund von KI zu einem Verhalten kommt, das wir als nicht richtig empfinden.

¹ TÜV = technischer Überwachungsverein = contrôle technique

Wie viel Verantwortung liegt bei den Unternehmen, die KI-basierte Algorithmen einsetzen?

Bei der Erstellung von Algorithmen darf man nicht systematisch auf die gute Intention des Einzelnen setzen. Das funktioniert nicht. Die Ethik muss in die Systeme eingeschrieben sein. Und langfristig wird es bei Unternehmen so sein, dass sie Strafzahlungen oder den Verlust ihrer Reputation riskieren, wenn sie bestimmte ethische Richtlinien bei ihren Algorithmen nicht einhalten.

Ihr Institut will Richtlinien darüber festlegen, was ethisch "richtig" ist. Definieren verschiedene Staaten das ethisch Wünschenswerte nicht unterschiedlich?

Klar: Wer an China denkt, denkt eher an Überwachungstechnologien als an den Schutz der Privatsphäre. Aber wir haben die ethischen Prinzipien miteinander verglichen, die in Europa, den USA und China zum Tragen kommen und eine überraschend hohe Übereinstimmung in bis zu 90 Prozent der Fälle gefunden. Insofern sind wir da nicht so weit voneinander entfernt.