
Ethik in der Robotik

Die Fortschritte auf den Gebieten Robotertechnik und künstliche Intelligenz führen dazu, dass in immer mehr Bereichen Robotersysteme eingesetzt werden. Das vorliegende Factsheet erläutert die damit einhergehenden ethischen Herausforderungen und stellt ein Instrument vor, mit dem die relevanten Fragen behandelt werden können.

Ein Roboter, was ist das?

Roboter sind überall, jedoch erkennen wir sie häufig nicht, da wir uns ein bestimmtes Bild von ihnen gemacht haben. Aufgrund von Filmen, Trickfilmen und Science-Fiction-Büchern stellen wir uns Roboter sehr menschenähnlich vor. Aber ein Industrieroboter, der am Fließband Metallteile zusammenfügt, hat keine kulturelle Präsenz. Für Robotikfachleute ist diese verbreitete Vorstellung von Robotern Fluch und Segen zugleich. Ein Segen, weil sich alle für Robotik interessieren, ein Fluch, weil die Erwartungen überhöht sind. Der reale Roboter ist fast dazu verdammt, zu enttäuschen.

Doch wovon sprechen wir denn eigentlich genau? Der amerikanische Robotiker Georges Bekey hat eine der bekanntesten Definitionen eines Roboters vorgeschlagen: eine Maschine, die sich in der Welt befindet, die wahrnimmt, denkt und handelt (*a machine, situated in the world, which senses, thinks and acts*).

- **Maschine:** Alle von Menschen geschaffenen Werkzeuge, unabhängig von Material, Form oder eingesetzter Mechanik.
- **In der Welt:** Der Roboter kann sich fortbewegen, greifen, stossen, ziehen und andere Aktionen ausführen und wirkt damit auf seine physische Umgebung ein. Diese Fähigkeit, direkt auf die Welt – und damit auf die Menschen – einzuwirken, macht Roboter zu einer eigenen, besonderen Spezies. Aufgrund dieser Eigenschaft gehören rein digitale Werkzeuge nicht dazu.
- **Die wahrnimmt, denkt und handelt:** Der Roboter wird geschaffen, um ein Ziel zu erreichen. Aus diesem Grund wird er befähigt, seine Umgebung wahrzunehmen. Diese Informationen verarbeitet er mithilfe von Funktionen, die üblicherweise mit Denken und Intelligenz assoziiert werden. Anschliessend wählt er Aktionen aus, die er in seiner Umgebung durchführt. Maschinen, die vollumfänglich ferngesteuert sind, und Automaten, die eine einfache Bewegung ausführen, gelten somit nicht als Roboter.

Evaluationsschema

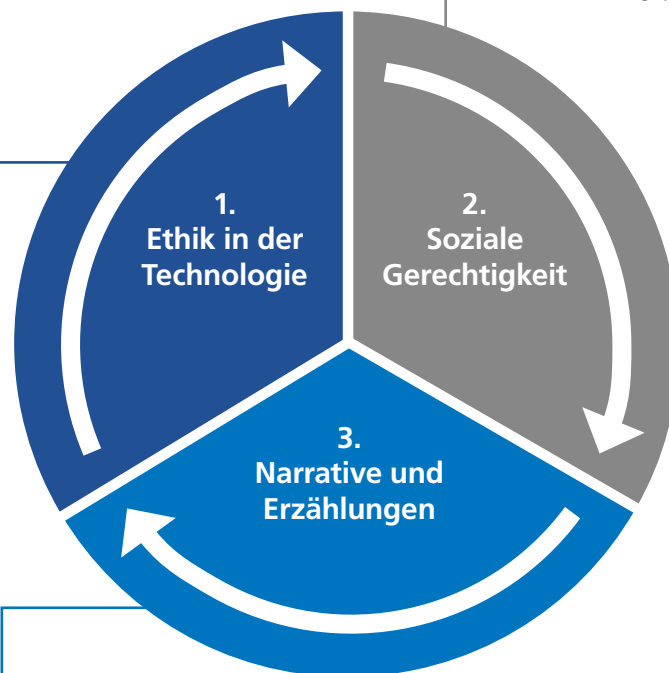
Dieses Evaluationsschema dient dazu, die zentralen ethischen Fragen bei einem Robotikprojekt zu strukturieren. Es eignet sich für Entwicklungsteams, politische Verantwortungstragende und alle, die sich für die Herausforderungen im Zusammenhang mit Technologie interessieren.

Ethik in der Technologie

- Ist der Roboter autonom und wenn ja, in welchem Sinn? Wer übernimmt bei Problemen die Verantwortung?
- Agiert der Roboter in einem Kontext, in dem er moralische Regeln einhalten muss? Wer hat bestimmt, welche Regeln das sind?
- Wenn Sie sich ein Katastrophenszenario vorstellen müssten, bei dem der Roboter zweckentfremdet wird, wie würde dieses aussehen?
- Könnte die äussere Erscheinung des Roboters und seines Interaktionssystems Nutzer:innen täuschen oder gewisse Stereotype verstärken?

Soziale Gerechtigkeit

- Verändert der Roboter die Interaktionen zwischen Menschen, zum Beispiel die Dynamik im Team?
- Kann der Roboter in einem Kontext, in dem menschliche Interaktion wichtig ist, eine Wirkung entfalten?
- Würden die Menschen, die mit dem Roboter interagieren sollen, angemessen konsultiert und geschult? Wurde gegebenenfalls ein alternatives Anforderungsprofil erstellt?



Narrative und Erzählungen

- Bedienen der Einsatz des Roboters und die damit einhergehende Kommunikation ein Narrativ der Konkurrenz zwischen Roboter und Mensch?
- Wird der Roboter als Beinahe-Mensch oder aber als Werkzeug im Dienste des Menschen präsentiert?

Eine Orientierungshilfe für Ethik in der Robotik

Roboter werfen zahlreiche ethische Fragen auf. Die Definition von Ethik ist hier weit gefasst und meint eine systematische Reflexion über Werte und deren Umsetzung in soziotechnischen Systemen, Systemen also, die sowohl eine technische als auch eine kontextuelle Komponente aufweisen. Die Orientierungshilfe behandelt drei grosse Themenbereiche und soll die relevantesten ethischen Fragen rund um Roboter einordnen. Sie ist gewissermassen eine geführte Tour für alle, die das Gebiet Ethik in der Robotik entdecken möchten. Im Zusammenhang mit verantwortungsvoller Innovation eignet sie sich insbesondere für Teams, die in die Entwicklung und die Konzeption von technologischen Geräten involviert sind. Die drei Themen bilden ein Ganzes und die Orientierungshilfe sollte iterativ eingesetzt werden.



Ethik in der Technologie

Das erste Thema sind ethische Fragen in der Technologie. Während der Erfindung, des Designprozesses und der Herstellung eines Roboters stellen sich denen, die den Roboter erschaffen, zahlreiche Fragen. Drei davon sind von besonderer Bedeutung: die Autonomie des Roboters, der moralische Roboter und das Design des Roboters.

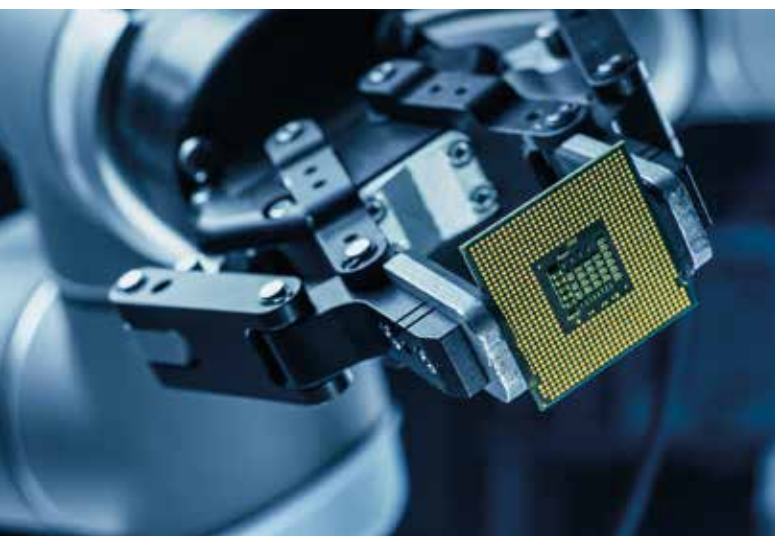
Erste Frage: Ist ein Roboter autonom? Die Autonomie des Roboters wird von Ethiker:innen und Ingenieur:innen unterschiedlich definiert. Die ISO-Normen zu Robotik liefern uns ein gutes Beispiel für die «Ingenieurvariante» des Autonomiekonzepts. Dort ist Autonomie definiert als «die Fähigkeit, bestimmte Aufgaben auszuführen, ausgehend vom bestehenden Zustand und von Erkenntnissen sowie ohne menschliches Eingreifen». Autonomie ist als Fähigkeit definiert, die Mittel zu bestimmen, um ein vorab festgelegtes Ziel zu erreichen. Die Fortschritte im Bereich KI steigern die Autonomie, da der Roboter bei der Auswahl und dem Einsatz der verwendeten Mittel immer besser wird. Eine grosse Herausforderung besteht darin, mit klaren Informationen zu den Verhaltensregeln einen Rahmen für den Roboter abzustecken, und so seine Fähigkeit, irgendeine beliebige Strategie auszuwählen, auf ein Ziel zu beschränken.

Die Fachleute der Ethik fügen dem Begriff der Autonomie eine weitere Ebene hinzu: Die Fähigkeit, sich Ziele zu setzen und diese

zu überdenken. Der Roboter hat diese Fähigkeit nicht, er führt aus, was andere ihm vorgegeben haben. Dieser erweiterte Autonomiebegriff wird oft in Verbindung mit dem Ichbewusstsein definiert. Dass wir uns unserer Existenz, unserer Bedürfnisse und unserer Werte bewusst sind, ermöglicht Autonomie. Der Roboter hat kein Bewusstsein seiner Existenz als Roboter.

Diese Unterscheidung zwischen zwei Ebenen der Autonomie spielt eine wichtige Rolle bei der Frage der Verantwortung. Wenn der Roboter nicht autonom ist, kann er nicht zur Verantwortung fähig sein. Der Roboter ist ein Akteur, er kann für gewisse Ereignisse verantwortlich sein, aber er ist kein moralischer Akteur. Ein Mensch muss immer verantwortungsvoll sein. Die verschiedenen Akteur:innen, die einen Roboter erfinden, konzipieren, produzieren und schliesslich nutzen, können unterschiedliche Verantwortung tragen. Entwickler:innen treffen eine bestimmte Grundsatzentscheidung darüber, was ein Roboter tun kann, den Herstellern des Roboters können Fabrikationsfehler unterlaufen, und die Nutzer:innen können ihren Kauf unsachgemäss verwenden. Mit fortgeschrittenen KI-Systemen ausgestattete Roboter, die ihr Verhalten ausgehend von vergangenen Erfahrungen verbessern können, stellen eine grosse Herausforderung dar, wenn es darum geht, Verantwortung zuzuweisen. Es ist schwierig, das Verhalten eines Roboters anhand der Rekonstruktion seiner Entstehung zu erklären. Trotzdem lässt sich die Diskussion über die Verantwortung in den gewohnten Kategorien führen. Solange die Entscheidungen der Entwickler:innen diejenigen des Roboters schon im Voraus bestimmen – auch die autonomen Entscheidungen –, kann der Roboter nicht als moralischer Akteur und somit nicht als verantwortungsfähig gelten.

Zweite Frage: Ist ein Roboter fähig, moralisch zu handeln? Die Frage der Autonomie kommt hier in anderer Form erneut ins Spiel. Ein Roboter kann moralisch handeln, sofern man ihn entsprechend programmiert. Aber er versteht nicht auf dieselbe Weise wie ein Mensch, was er tut. Er führt einfach die Programmierung aus. Programmiererteams stehen somit vor einer dreifachen Herausforderung: Welchen ethischen Ansatz sollen sie wählen, beispielsweise eine an den Konsequenzen oder aber eine an der Pflichterfüllung orientierte Ethik? Und wie sollen die zu program-



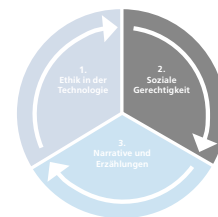


mierenden Regeln, über die Konsens herrscht, definiert werden? Diese Diskussion kann nur in Abhängigkeit vom Kontext erfolgen, wobei es zum Beispiel darauf ankommt, ob es sich um Pflegeroboter oder autonome Waffensysteme handelt. Und wie kann man gewährleisten, dass der Roboter technisch dazu in der Lage ist, diese Regeln einzuhalten? Wenn der Roboter mögliche Konsequenzen erkennen und bewerten und anschliessend entsprechend handeln muss, wie kann man sicher sein, dass er sich dabei auf die richtigen Informationen stützt? Dies ist eine zentrale Herausforderung, für die es momentan noch keine Lösung gibt.

Dritte Frage: Wie soll das Design eines Roboters aussehen? Wird ein Roboter konzipiert, gilt es eine Form und eine äusserer Erscheinung zu wählen, aber auch eine Stimme und ein Nutzerinterface. Diese Notwendigkeit, eine physische Hülle für den Roboter zu gestalten, unterscheidet die Maschine von digitalen Einheiten. Dieses äussere Design beeinflusst das Vertrauen, das wir Menschen in diesen Roboter haben. Menschen möchten mit Einheiten interagieren, die sie (wieder-)erkennen und die sie in einem spezifischen Kontext als kohärent erleben. Eine ungewohnte Kombination aus Elementen, die als menschlich wahrgenommen werden (Formen, Texturen, Verhaltensweisen), und Elementen, die mit der Welt der Roboter assoziiert werden, kann befremdend oder etwas unheimlich wirken. Somit stellt sich beim Design eines Roboters die Frage nach der Manipulation. Ist es wünschenswert, einen Roboter herzustellen, der einem Menschen oder einem Tier gleicht, um den Nutzenden etwas vorzutäuschen? Gleich wie bei digitalen Interaktionen – beispielsweise mit einem Chatbot – scheint es nötig, klare Informationen darüber zu haben, wer genau an einer Interaktion beteiligt ist. Neben dem Aspekt der Manipulation wirft das Design eines Roboters auch Fragen bezüglich Repräsentation auf. Wählt man für den Roboter gewisse charakteristische Eigenschaften, wird ein bestimmtes, allgemeingültiges Bild damit projiziert. So haben die meisten Sexroboter grosse Brüste, schmale Hüften und ein kindliches Gesicht. Solche Designentscheidungen können Stereotype verstärken. Die Robotik kann uns allerdings auch ermöglichen, bestimmte Fragen zu erkunden, zu denen Menschen normalerweise keinen Zugang haben: Muss ein Roboter ein erkennbares Geschlecht haben, oder ist ein möglichst neutraler Roboter vorzu-

ziehen? Für einige kritische Stimmen könnte Robotik eine Gelegenheit sein, fließendere Übergänge zwischen den Geschlechtern auszuprobieren und weniger abhängig zu sein von der Binarität Mann/Frau, die unser soziales Leben prägt.

Diese drei Fragen schlagen eine Brücke zu einer der wichtigsten Erkenntnisse der Technikethik: Technologische Innovation ist nie neutral. Sie wird mit der expliziten oder impliziten Absicht erarbeitet, entwickelt und hergestellt, bestimmten Zielen und Werten zu entsprechen. Alle damit zusammenhängenden Entscheidungen werden als ethische Fragen betrachtet, denn sie hinterfragen bestimmte Werte und auferlegen den Entwicklungs- und Produktionsteams Verantwortung.



Soziale Gerechtigkeit

Das zweite Thema betrifft soziale Gerechtigkeit im Zusammenhang mit dem Einsatz von Robotern. Es geht darum, wie sich Roboter, und die Robotik allgemein, auf die Gesellschaft auswirken. Dabei ist zu analysieren, was Roboter für menschliche Beziehungen und den Arbeitsmarkt bedeuten. Beim kontextuellen Ansatz stehen die verschiedenen Anwendungsgebiete wie Pflege, Bildung oder Sicherheit im Fokus.

Der Einfluss auf den Arbeitsmarkt gilt als eine der bedeutendsten sozialen Konsequenzen der Robotik. Ein Vorteil von Robotern ist, dass sie die Menschen von gefährlichen, repetitiven oder stumpfsinnigen Aufgaben befreien. Die Kehrseite der Medaille: Robotern wird zugleich nachgesagt, dass sie Arbeitsplätze vernichten. Die Modelle, die Wirtschaftsfachleute entwickelt haben, um diese Entwicklung zu beziffern, umfassen mehrere Aspekte. Einerseits hängt es von der Art des Wirtschaftsbereichs ab, wie stark Arbeitsplätze durch Robotisierung gefährdet sind. In Branchen, die

auf planbaren repetitiven Prozessen beruhen und wenig menschliche Interaktion erfordern, ist die Gefahr des Verlusts von Arbeitsplätzen am grössten. Andererseits trifft die Robotisierung nicht alle Arbeitnehmenden im gleichen Ausmass. Am ehesten laufen Geringqualifizierte in Industriesektoren Gefahr, durch Roboter – im Gegensatz zu KI – ihre Arbeit zu verlieren. Bei Arbeitnehmenden, die kollaborative Roboter in ihren Alltag integrieren müssen, wird sich das Pflichtenheft erweitern. Die Auswirkungen der Robotisierung reichen aber über die Interaktion zwischen Mensch und Roboter hinaus. Die Integration eines Roboters verändert die Beziehungsdynamik zwischen Menschen, zum Beispiel innerhalb eines Teams. Da der Roboter auf Strukturierung und Messung angewiesen ist, trägt er dazu bei, Verhalten zu normieren. Ein Roboter kann nicht improvisieren. Damit er funktioniert, muss er eingegrenzt werden, und diese Eingrenzung wirkt sich wiederum auf das Personal und die Institution aus. Den Überlegungen zur sozialen Gerechtigkeit liegt

die grosse Frage nach der Beziehung zwischen Anstellung, Arbeit, Existenzgrundlage und sozialer Anerkennung zugrunde. Drei Aspekte können bereits heute umgesetzt werden (und sind es teilweise bereits). Erstens müssen wir die Art und Weise anpassen, wie wir das ganze Leben lang lernen. Zweitens braucht es eine sozialpolitische Antwort auf die Veränderungen, die auf dem Arbeitsmarkt im Gange sind. Wird ein Wirtschaftsbereich von einer schnellen und tiefgreifenden technologischen Veränderung getroffen, muss die Gesellschaft den Schock für die direkt Betroffenen abfedern können. Die Idee des verantwortungsvollen Wandels ist entscheidend: Die technologischen Veränderungen müssen als permanenter Wandel betrachtet werden und nicht als Katastrophe, die bereits ihren Lauf nimmt und aus der es einen Ausweg zu finden gilt. Drittens müssen wir darüber diskutieren, welche Bereiche von der Robotisierung ausgenommen sein sollten. Dieser Ansatz zielt darauf ab, bestimmte Lebenssituationen zu identifizieren, in denen ein Kontakt zu einem anderen Menschen erhalten bleiben soll.



Narrative und Erzählungen

Das dritte Thema beinhaltet die grossen gesellschaftlichen Erzählungen rund um die Robotik. Diese Narrative bilden den allgemeinen Rahmen unserer Diskussionen über Roboter und sind zugleich eine kollektive Geschichte, die wir pflegen. Im Bereich Robotik strukturieren drei Narrative einen Grossteil der gesellschaftlichen Debatte. Hier soll die ethische Dimension dieser Narrative und der kollektiven Vorstellungswelt aufgezeigt werden. Sprache ist nicht neutral und sie trägt eine bestimmte Sicht der Realität in sich. Zudem haben bestimmte Akteur:innen ein wirtschaftliches oder politisches Interesse daran, die Realität aus einer spezifischen Perspektive darzustellen.

Beim ersten Narrativ geht es um die allgemeine Beziehung zwischen Robotern und Menschen. Diese oszilliert zwischen Konkurrenzkampf und Kooperation. In der Science-Fiction mündet der Konkurrenzkampf früher oder später in eine Situation, in der die Roboter die Menschen zu kontrollieren oder zu unterwerfen drohen. Im Gegensatz dazu stellen wir uns die Beziehungen zwischen Robotern und Menschen aber auch als partnerschaftliches Verhältnis vor. Der Roboter ist den Menschen zu Diensten, er ist vergleichbar mit einem perfektionierten Werkzeug, das die Menschen einsetzen, um ihre Ziele zu er-

reichen. Welches Narrativ für das Verhältnis zwischen Roboter und Mensch gewählt wird, beeinflusst tiefgreifend, wie der Roboter wahrgenommen wird (als mehr oder weniger komplexes Ding oder als neue Art von Person).

Dieses erste grosse Narrativ zwischen Konkurrenz und Kooperation führt zum zweiten grossen Narrativ, demjenigen über menschliche Eigenschaften. Die Spiele Schach und Go, die Kreativität, Einfallsreichtum und Taktik erfordern, galten lange Zeit als dem Menschen vorbehalten. Heute muss sich der Mensch dabei geschlagen geben. Müssen wir weiter danach suchen, was den Menschen ausmacht, oder im Gegenteil den Gedanken akzeptieren, dass der Mensch einfach ein Lebewesen wie alle anderen ist? Hier spielt auch die Diskussion über die Unterschiede zwischen Mensch und Tier sowie deren Platz in der Natur mit hinein.

Diese Suche nach den menschlichen Eigenschaften bringt uns zum dritten grossen Narrativ, dem über den Sinn des menschlichen Lebens. Roboter werden oft als Höhepunkt einer arbeitsamen Gesellschaft dargestellt, die voll und ganz auf Produktivität ausgerichtet ist. Sie schufteten ohne Unterbruch. Die Robotisierung stellt die Bedeutung der entschädigten Arbeit (Erwerbsarbeit), als wichtiger Pfeiler der persönlichen Identität, grundsätzlich infrage. Wenn unsere Grosskinder einst vier Stunden pro Woche in einer grösstenteils robotisierten Gesellschaft arbeiten, worin besteht dann der Sinn ihres Lebens? Welche Narrative können wir erschaffen, um diese Gesellschaft zu begleiten? Die Frage nach dem Sinn des Lebens steht ganz oben auf der Agenda.

Die SATW fördert Projekte im Bereich Künstliche Intelligenz

Zwei wichtige Entwicklungen verändern unsere Beziehung zu Robotern: die zunehmende Interaktion der Roboter mit der breiten Öffentlichkeit und die Verbesserung der künstlichen Intelligenz, die in Robotern eingebaut ist. Diese tiefgreifenden Veränderungen konfrontieren uns mit einer Reihe von ethischen Herausforderungen, die wir als Individuen und als Gesellschaft lösen müssen. Es geht darum, mit welchen Arten von Robotern wir interagieren wollen, wie diese programmiert werden und wie wir die Gesellschaft auf die Mensch-Roboter-Interaktionen vorbereiten, um eine bessere Lebensqualität zu erreichen.

Diese Fragen sind von äusserster Wichtigkeit. Dennoch findet die Debatte häufig nicht statt und wenn sie das doch tut, wird sie mit Klischees und extremen Szenarien geführt. Die Ausbildungssituation für zukünftige Ingenieur:innen ist ebenfalls nicht zufriedenstellend. Das Projekt «Ethik in der Robotik» will eine dieser Herausforderungen angehen, indem es der breiten Öffentlichkeit die Werkzeuge für eine qualitativ hochwertige Debatte an die Hand gibt.

Die SATW unterstützte ein Team von EPFL-Forschenden unter der Leitung des SATW-Mitglieds Prof. Francesco Mondada in der Planung, Umsetzung und Bekanntmachung des Pilotprojekts, das in diesem Factsheet vorgestellt wird.

Für den Bereich Künstliche Intelligenz (KI) verfügt die SATW über eine Themenplattform und ein Schwerpunktprogramm; sie organisiert regelmässig Dialogformate und gibt Publikationen heraus – beides in Zusammenarbeit mit Expert:innen wie Partnerorganisationen. Die SATW ist auch im Bereich Daten sehr aktiv, denn ohne Daten funktioniert KI nicht. Im Frühling 2021 hat die SATW gemeinsam mit der Direktion für Völkerrecht des EDA, dem Bundesamt für Kommunikation und der Swiss Data Alliance das nationale «Netzwerk Digitale Selbstbestimmung» gegründet, um Bürger:innen, Unternehmen und öffentlichen Einrichtungen eine Nutzung der Datenwirtschaft auf der Basis freiheitlich-demokratischer Grundwerte zu ermöglichen.

Impressum

Autor: Johan Rochel (EPFL, UZH, Ethix)

Projektleiter: Francesco Mondada (EPFL)

Projektteam: Johan Rochel (EPFL, UZH, Ethix), Alexandre Luyet (SATW), Adriana Cantaluppi (SATW)

Redaktion: Esther Lombardini

Übersetzung: weiss traductions genossenschaft

Grafik: Andy Braun

Mai 2022



Bei EPFL Press erhältlich:

Les robots parmi nous: pour une éthique des machines

satw it's all about
technology

Schweizerische Akademie der Technischen Wissenschaften SATW
St. Annagasse 18 | 8001 Zürich | 044 226 50 11 | info@satw.ch | www.satw.ch

Fotos: Adobe Stock

