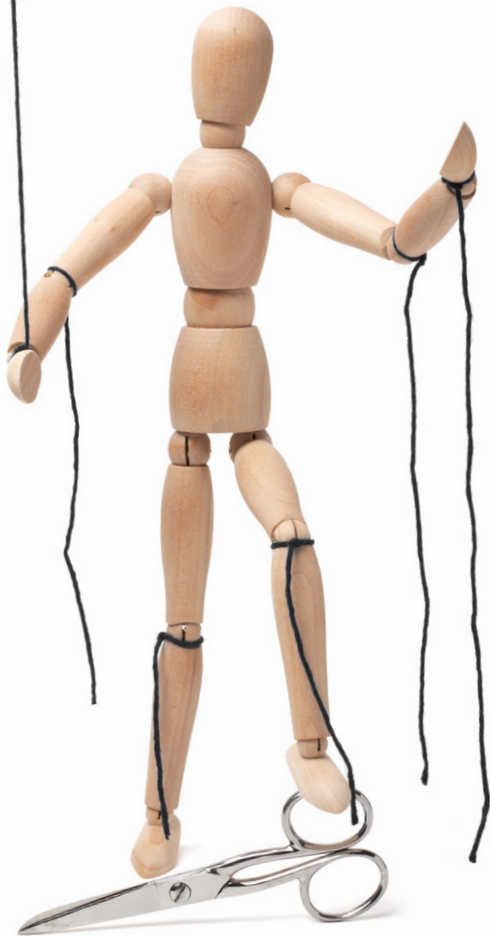


ethix

Innovation Brief #01



Autonome  
Systeme

Worum es geht →

Was ist ein autonomes System –  
und was hat es mit Innovationsethik zu tun?

Zukunftsszenarien & Branchen →

Autonome Systeme in der Welt von morgen.

Ethische Risikogebiete →

Verantwortung, Sicherheit, Kontrolle und Entscheidung.

Fokus →

Autonomie und Verantwortung.

Blick nach vorne →

Chancen für die Zukunft von autonomen Systemen.

ethix Ressourcen →

Mit ethix Chancen und Risiken entdecken.





Selbstfahrende Autos,  
die uns vom Bahnhof abholen.

Algorithmen, welche die perfekte  
Lieferzeit berechnen und Paketdrohnen  
auf den Weg schicken.

Sich selbst füllende Kühlschränke,  
die unsere Vorlieben lernen und  
selbständig neue Süßigkeiten bestellen.

Vollständig autonome Heimelektronik und  
ausser Kontrolle geratene Roboterkriege.

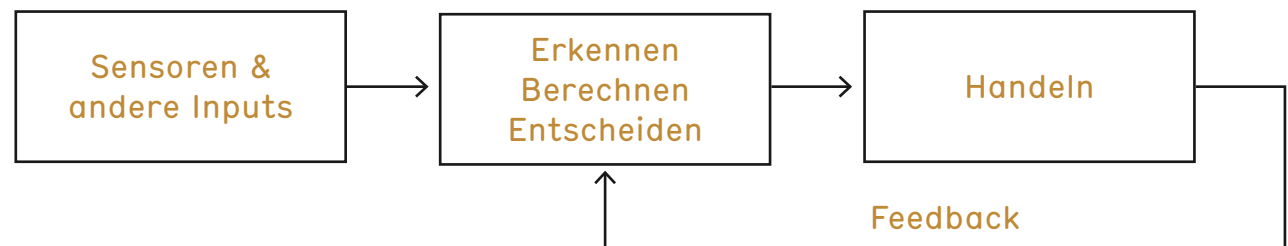
So könnte die Welt der Zukunft aussehen,  
die von autonomen Systemen bevölkert ist.  
Was früher jenseits der Grenze des  
Möglichen schien, ist heute ein konkret  
diskutiertes Zukunftsszenario.

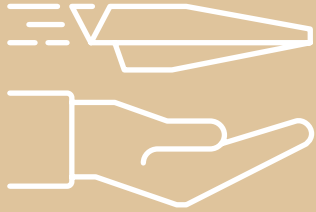
## Worum es geht

### Was ist ein autonomes System?

Menschen sind autonome Systeme. Wir nehmen die Welt um uns herum wahr, überlegen uns, wie wir vorgehen müssen, um unsere Ziele zu erreichen und handeln dann entsprechend. Autonome Systeme wie selbstfahrende Autos sind ganz ähnlich. Sie nehmen mit Sensoren und anderen Inputs die Umgebung wahr. Darin erkennen sie Gegenstände, berechnen verschiedene Handlungsmöglichkeiten und entscheiden sich für eine davon. Als Basis dienen ihnen die Wahrnehmung und vorprogrammierte oder selbst gelernte Ziele. Am Ende nehmen sie selbständig auf die Umgebung Einfluss, indem sie zum Beispiel von A nach B fahren. Ein wirklich autonomes System kann sich also ohne Hilfe von jemand anderem in der Welt bewegen und handeln.

Wie gelangt ein autonomes System zu einer Entscheidung?





## Betroffene Branchen

Robotik  
Drohnen  
Fahrzeuge  
Dual-Use  
Sensorik  
Maschinelles Lernen  
Mobilität  
Infrastruktur  
Öffentlicher Verkehr  
Privatverkehr  
Logistik  
Lagermanagement  
Autonome Fahrzeuge  
Drohnenlieferungen  
Pflege  
Pflegeroboter  
Automatisierte Diagnosen

## Zukunftsszenarien

Die Firma Oberholzer produziert Motoren für elektrische Zahnbürsten.

Ein Rüstungshersteller möchte diese Motoren gerne kaufen, da sie wie geschaffen sind für die Feinsteuerung eines automatischen Maschinengewehrs. Die Firma geht nach einer Besprechung mit den Mitarbeitenden nicht auf die Anfrage ein. Tut sie das einzig Richtige?

Das Startup Mapfly produziert autonome Drohnen, die selbstständig ein Gebiet abfliegen und kartografieren können.

Nach einer Anfrage liefert es hundert Drohnen an ein Land, in dem ein Bürgerkrieg tobt.

Lassen sich solche Lieferungen rechtfertigen?

Ein Automobilhersteller hat einen selbstfahrenden Lastwagen entworfen, der innerhalb von 5 Jahren die Lohnkosten einer Fahrerin amortisiert.

Nach 10 Jahren gibt es in den USA und ersten europäischen Ländern nur noch selbstfahrende Lastwagen.

Wer ist dafür verantwortlich, dass die betroffenen Mitarbeitenden anderswo eine qualifizierte Arbeit finden können?





Testen Sie Ihre Risikogebiete mit dem ethix Fragebogen.

# Ethische Risikogebiete

## Die Frage der Verantwortung

Wer ist für die Entscheidungen eines selbstfahrenden Fahrzeugs verantwortlich? Die Fahrerin, die Herstellungsfirma, der Anbieter eines Mobilitätsservices, die Programmiererin, die Zulassungsstelle oder vielleicht sogar das Fahrzeug selbst?

## Die Pflicht zur Sicherheit

Autonome Systeme könnten den Strassenverkehr, die Flugsicherung oder andere Lebensbereiche sicherer machen. Gleichzeitig stellt sich die Frage, ob wir Entscheidungen, bei denen Menschenleben auf dem Spiel stehen, einem emotionslosen Algorithmus überlassen möchten. Und vor allem müssen die Sensoren und Algorithmen in allen möglichen klimatischen und sozialen Umständen funktionieren.

## Das Trolley Problem

Ist es ethisch legitim, eine Person aktiv zu schädigen, um eine Gruppe anderer zu schützen? Darin wird die weitreichende Uneinigkeit zwischen verschiedenen ethischen Prinzipien sichtbar. Die Programmierung eines Entscheidungs- und allfälligen Lernalgorithmus hängt stark davon ab, für welche ethischen Prinzipien wir uns entscheiden. Wie können wir Maschinen programmieren, die regelmässig gute Entscheidungen fällen?

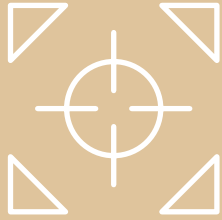
## Dual-Use

Armee und Rüstungsfirmen haben ein grosses Interesse an autonomen Systemen. Aber auch viele zivile Technologien wie Software, Elektronik oder Drohnen finden regelmässig den Weg in Kriegsgebiete. Produzentinnen und Produzenten müssen sich hier schwerwiegenden Entscheidungen stellen.

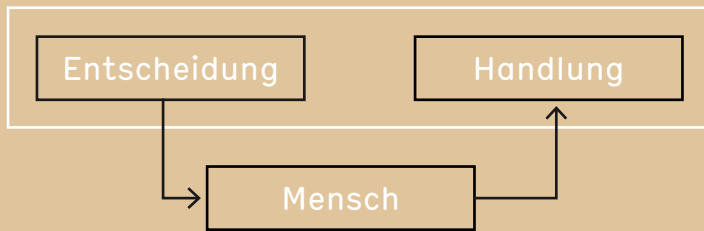
## Negative Nebeneffekte

Ohne sorgfältige Planung könnte maschinelles Lernen schnell zu unerwünschten und unerwarteten Nebeneffekten führen. Die Möglichkeiten reichen von der vollständigen Automatisierung einer Branche über ein unaufhaltsames enthumanisiertes Wettrüsten und ausser Kontrolle geratener Roboterkräfte bis zur Auslöschung der Menschheit.

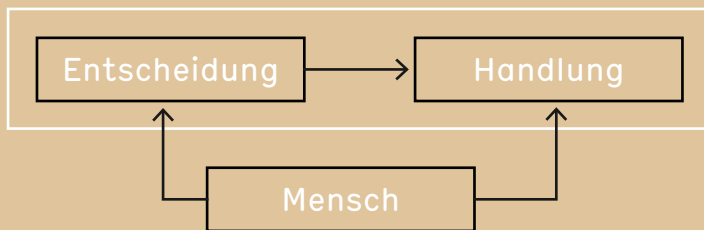




Bei autonomen Systemen mit «Human in the Loop» Strategie löst ein Mensch die Handlung aus.



Bei autonomen Systemen mit «Human on the Loop» Strategie beobachtet ein Mensch und greift notfalls ein.



## Fokus

Wer ist verantwortlich, wenn ein selbst-fahrendes Auto verunfallt? Wollen wir autonomen Waffensystemen die Entscheidung über Leben und Tod überlassen? Und wie sollen diese Systeme in ethisch aufgeladenen Situationen entscheiden?

Die Frage der Verantwortung sticht bei autonomen Systemen besonders hervor. Denn wer autonom entscheiden kann, ist grundsätzlich auch für die Entscheidung verantwortlich.

Bei programmierten Maschinen ist dieser Zusammenhang besonders prekär. Denn eine ProgrammiererIn kann nicht jede Situation bereits im Voraus antizipieren. Das Resultat einer Entscheidung ist darum in den meisten Fällen nicht vorhersehbar. Deshalb braucht es gut ausgewählte ethische Prinzipien, an denen sich das autonome System orientiert.

Doch wer eine moralische Maschine programmieren möchte, fällt automatisch selbst schwerwiegende ethische Entscheidungen. Die bisherige Forschung dazu ist noch wenig ausgereift. Manche versuchen herauszufinden, wie sich Menschen in ethischen Dilemmas verhalten, um empirisch Prinzipien zu generieren. Doch die Ethikkommission des Bundesministeriums für Verkehr und digitale Infrastruktur kommt zum Schluss, dass «rechte dilemmatische Entscheidungen nicht

ethisch zweifelsfrei programmierbar» seien. Weil eine solche moralische Maschine noch nicht existiert, sind die meisten autonomen Systeme bisher menschlicher Kontrolle unterworfen. Bewaffnete Drohnen können zum Beispiel weitgehend selbstständig starten und landen, Landstriche absuchen und Ziele identifizieren. Doch noch immer ist ein Mensch am Hebel, der die folgenschwere Entscheidung treffen muss. Das nennt sich «Human in the loop», da sich ein Mensch in die Entscheidungsschleife einmischt.

Andere Systeme gehen einen Schritt weiter in Richtung kompletter Autonomie und geben dem Menschen bloss noch eine Überwachungsfunktion. Diese «Human on the loop» Strategie findet bei vielen selbstfahrenden Autos Anwendung, die zwar bereits selbstständig Entscheidungen treffen können, aber immer noch von der Fahrerin überwacht werden.





Wie möchten wir die Welt mit autonomen Systemen teilen? Wie bei jeder Technologie gibt es Lichtblicke und Schattenseiten. Wenn Maschinen und Algorithmen selbstständig Entscheidungen fällen, ist es besonders wichtig, dass wir die Kontrolle nicht verlieren und verstehen können, wie eine spezifische Entscheidung entsteht. Doch wir sollten auch das Potenzial nutzen, um einen tatsächlichen Fortschritt zu erzielen.

## Blick nach vorne

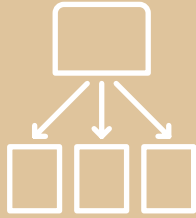
Autonome Systeme eignen sich besonders gut dafür, Abläufe **einfacher, sicherer und effizienter** zu gestalten. Innovation sollte sich auf diese Bereiche fokussieren und neue Möglichkeiten aufzeigen, um das Leben von Menschen zu verbessern und zu erleichtern.

Zum Beispiel könnten lernende Algorithmen Zusammenhänge entdecken, die uns Menschen verborgen bleiben. Dadurch können sie die Aufmerksamkeit und menschlichen Ressourcen auf jene Bereiche lenken, in denen wir mit unseren einzigartigen Fähigkeiten zum Zug kommen. So können autonome Systeme **unterstützend wirken**, Daten strukturieren und eine verbesserte Übersicht in komplexen Bereichen schaffen.

Auch im Bereich der Sicherheit sind autonome Systeme hilfreich und Menschen oft überlegen. Denn sie machen weniger Fehler und sind besser darin, gefährliche Routine- oder Fließbandarbeit zu erledigen. Dadurch können **Menschen in gefährlichen Arbeitsbereichen entlastet und geschützt** werden.

Schliesslich haben lernende Systeme das Potenzial, **beschränkte Ressourcen besonders effizient** einzusetzen und zu verteilen. Sei es im Bereich von Smart Grids und der Stromproduktion oder überall sonst, wo knappe Ressourcen verwendet werden. Mit autonomen Systemen können unnötige Verluste erkannt und verhindert werden.





Mittels Instrumenten von ethix und anderen Partnern können schon während der Entwicklungsphase und des Designprozesses Chancen und Risiken von autonomen Systemen analysiert und diese gegebenenfalls verbessert werden.

## Ressourcen

### ethix Mapping und ethix Canvas

Zur ersten Abklärung von Verantwortungsfragen und ethischen Risikogebieten von Innovationen im Bereich autonome Systeme.

### ethix Workshop

Zur vertieften Auseinandersetzung mit ethischen Risiken und Chancen autonomer Systeme und eventueller Implementierung von Checklisten und anderen Werkzeugen im Innovationsprozess.

### Interne Schulungen

Schulung und Sensibilisierung der am Innovationsprozess eines autonomen Systems beteiligten Mitarbeitenden für ethische Risikogebiete und Chancen.

### Labels

Beispielsweise Anwendung eines bestehenden Labels für verantwortungsvolle autonome Systeme.

Begleitung der Umsetzung solcher Normen im Unternehmensalltag durch ethix.

