



Maria Pawelec, Cora Bieß und Alexander Orlowski

Ethisch und sozial wünschenswerte Technikgovernance fördern



Points to Consider zu Technikgovernance

Handlungsempfehlungen zum Autonomen Fahren und zu Deepfakes

MATERIALIEN ZUR ETHIK IN DEN WISSENSCHAFTEN

BAND 17

herausgegeben vom
Internationalen Zentrum für Ethik in den Wissenschaften (IZEW)
Eberhard Karls Universität Tübingen

gefördert durch das
Bundesministerium für Bildung und Forschung



Diese Broschüre entstand im Rahmen des Forschungsprojekts TANGO (Technik-Innovation: sozialwissenschaftliche und ethische Analysen zu Governance)

unter der Projektleitung von
Prof. Dr. Regina Ammicht Quinn

verfasst von
Maria Pawelec, Cora Bieß und Alexander Orlowski

unter der Mitwirkung von
Jovanna Horn, Miriam Schröder und Anna Tilling

Vorgeschlagene Zitierweise:

Pawelec, Maria, Cora Bieß und Alexander Orlowski (2021): Ethisch und sozial wünschenswerte Technikgovernance fördern - Points to Consider zu Technikgovernance; Handlungsempfehlungen zum Autonomen Fahren und zu Deepfakes. Tübingen: IZEW, Materialien zur Ethik in den Wissenschaften, Band 17.

ISBN: 978-3-935933-19-3

Inhaltsverzeichnis

<i>I. Das Forschungsprojekt TANGO</i>	7
<i>II. Ethisch wünschenswerte Technikgovernance: Points to Consider</i>	11
a) Relevanz des Entwicklungskontexts für Governance	14
b) Berücksichtigung aller möglichen Regulierungskontexte	18
c) Chancen nutzen und bestehende Strukturen hinterfragen	20
d) Mündige Technologienutzung durch Aufklärung und Bildung	22
<i>III. Handlungsempfehlungen Autonomes Fahren</i>	26
<i>IV. Handlungsempfehlungen Deepfakes</i>	34

I. Das Forschungsprojekt TANGO

Diese Broschüre fasst die zentralen Ergebnisse des Projekts TANGO („Technik-Innovation: sozialwissenschaftliche und ethische Analysen zu Governance“) zusammen, das vom Bundesministerium für Bildung und Forschung (BMBF) von Januar 2019 bis Dezember 2020 gefördert wurde. TANGO war Teil der ITA-Förderlinie („Innovations- und Technikanalyse“).

Ziel des Projekts war es, aus einer Analyse unterschiedlicher Innovationsfelder abzuleiten, wie die Governance technologischer Innovationen erfolgen kann, um eine ethisch und gesellschaftlich wünschenswerte Technikentwicklung (und -anwendung) zu erreichen. Dazu kombinierte TANGO eine ethische Technikfolgenabschätzung in den beiden Innovationsfeldern mit der Analyse verschiedener Governancemechanismen, von (supra-)nationaler Regulierung über technische Normen bis hin zur Berufsethik. Die Forschungsarbeit im Projekt TANGO zeigte, dass in den Innovationsfeldern **autonomes Fahren** und **Deepfakes** zahlreiche ethische Herausforderungen stecken und deren Governance daher von großer gesellschaftlicher Bedeutung ist. Gleichzeitig berücksichtigen Governancemechanismen auf verschiedenen Ebenen unterschiedlich stark ethische, rechtliche und gesellschaftliche Aspekte, und variieren in ihrer Wirksamkeit innerhalb und zwischen den Fallstudien.

Vergleicht man die beiden Fallstudien miteinander, werden sowohl Unterschiede als auch Gemeinsamkeiten in der Governance der beiden Technologien deutlich. Auf Basis dieses Vergleichs formulierte TANGO Rückschlüsse auf die Wirksamkeit und Wünschbarkeit verschiedener Governancemechanismen für zukünftige Technikgovernance.

Daraus folgen vier generelle „Points to Consider“ (zu beachtende Punkte) in Bezug auf Innovationsgovernance, die sich den folgenden Überschriften zuordnen lassen:

- 1) Relevanz des Entwicklungskontexts für Governance
- 2) Berücksichtigung aller möglichen Regulierungskontexte
- 3) Chancen nutzen und bestehende Regulierungen hinterfragen
- 4) Mündige Technologienutzung durch Aufklärung und Bildung

Diese werden in der Broschüre näher ausgeführt. Darüber hinaus finden Sie im Anschluss daran spezifische Handlungsempfehlungen für die beiden Innovationsfelder autonomes Fahren und

Deepfakes, die in Zukunft eine ethisch und gesellschaftlich wünschenswertere Governance dieser Felder ermöglichen können.

Lesehinweis:

Die Punkte, die generell für die Governance von Innovation beachtet werden sollten, sind durch graue Hinterlegung hervorgehoben.



II. Ethisch wünschenswerte Technikgovernance: Points to Consider

Die aus dem Projekt abgeleiteten Handlungsempfehlungen für die Governance von Innovation sind in den hier vorliegenden Points to Consider zusammengefasst. Das Format der Points to Consider gibt keine einfachen Antworten auf komplexe Fragen, sondern regt normative Reflexionsprozesse an. Insofern sind die Projektergebnisse in unterschiedlichen Anwendungs- und Entwicklungskontexten verwertbar. Sie bieten Orientierungspunkte für eine gesellschaftlich wünschenswerten Technikentwicklung.

Die beiden **Technologien** autonomes Fahren und Deepfakes unterscheiden sich zwar fundamental in ihrer Entwicklungsumgebung und in ihren Anwendungskontexten, weisen jedoch auch zentrale **Überschneidungspunkte** auf: Es handelt sich bei beiden Technologien um KI-gestützte Entwicklungen. Eine Erkennung von manipuliertem Bildmaterial, wie durch Deepfake-Detektionsalgorithmen, ist auch für die Sicherheit des autonomen Fahrens zentral. Hier können Videoinformationen während der Übertragung an einen zentralen Rechner manipuliert werden, so dass die (externe) Bilderkennung die Situation auf der Straße falsch erkennt und nicht die richtigen Steuerungsbefehle übermittelt. Dadurch kann es zu Unfällen kommen, etwa wenn ein Stoppschild nicht als Stoppschild, sondern als Vorfahrtsschild erkannt werden würde. Deepfake-Detektionssysteme könnten somit für die Sicherheit des autonomen Fahrens in Zukunft eine Rolle spielen. Darüber hinaus kann die GAN-Technologie, auf der Deepfakes unter anderem basieren, genutzt werden um Fahrsituationen (etwa unterschiedliche Wetterverhältnisse) zu simulieren und anhand dessen Systeme des autonomen Fahrens zu trainieren und zu optimieren.

Es gibt jedoch auch große und entscheidende **Unterschiede** zwischen den beiden Technologien des autonomen Fahrens und der Deepfakes. So ist der **Anwendungskontext** des autonomen Fahrens klar definiert und im Vergleich zu Deepfakes weniger vielfältig und divers. Insgesamt sind **beim autonomen Fahren meist bestimmte Aspekte der Technologie** und ihrer Auswirkungen **umstritten** und ethisch problematisch. Bei **Deepfakes** hingegen ist der **Anwendungskontext der Technologie** viel zentraler: Ein und dieselbe Technologie etwa des Faceswap oder Facial Reenactment kann ebenso in einer Aufklärungskampagne gegen Kinderarbeit eingesetzt werden wie in Rachepornografie. Dies erschwert die ethische Bewertung von Deepfakes, die – wie die meisten KI-Anwendungen – ausschließlich kontextbezogen erfolgen kann. Gleichzeitig hat dies entscheidende Auswirkungen auf die Governance dieser Technologie, denn dieselben Governance-Mechanismen können je nach gesellschaftlicher Position

unterschiedliche Auswirkungen für Beteiligte und Betroffene hervorrufen. Governance von Technikinnovation sollte daher stets von **Abwägungsprozessen** zwischen **verschiedenen Kontexten** geprägt sein. Dabei ist es wichtig, dass in diesen Abwägungsprozess auch verschiedene Interessen mit einbezogen werden.

Governance-Überlegungen sind für Innovationen mit klar abgegrenztem Anwendungsbereich einfacher auszugestalten als für **Innovationen mit sehr breitem und diffusem Anwendungsbereich** wie etwa viele multifunktional einsetzbare KI-Innovationen. Hier müssen stets die **Auswirkungen verschiedener Governanceinstrumente in verschiedenen Kontexten gegeneinander abgewogen** werden.

Darüber hinaus gibt es große Unterschiede im **Entwicklungskontext** zwischen den beiden Feldern autonomes Fahren und Deepfakes. Im **autonomen Fahren** bündelt sich die Entwicklung bei **wenigen, großen Akteur*innen**. Zwar können auch kleinere, unabhängige Entwickler*innen zu dieser technologischen Innovation beitragen. Jedoch spielen gerade im deutschen Entwicklungskontext die Automobilhersteller wie Volkswagen, Daimler oder BMW die zentrale Rolle. In diesen Konzernen findet ein Großteil der Entwicklung statt. Selbst wenn die Technologie von kleineren Akteur*innen entwickelt wurde, wird sie in der Regel von den großen Automobilherstellern in ein Gesamtfahrzeug implementiert. In diesem Rahmen durchlaufen entwickelte Komponenten **Testverfahren**, bevor sie in Fahrzeuge eingebaut werden dürfen und auf der Straße zugelassen werden. Darüber hinaus gibt es technische Prüforganisationen wie den TÜV, die als Gatekeeper fungieren: Der TÜV überprüft das finale Produkt und entscheidet darüber, welche Technik letztendlich zugelassen wird. Zudem begleitet er den gesamten Entwicklungsprozess, sodass die Entwicklung kontinuierlich anhand von technischen Normen überprüft wird. Es werden verschiedene Umstände und Szenarien durchgespielt, in denen die Technik funktionieren sollte. Dadurch spielt die Absicherung der Entwicklung und damit des Produkts eine große Rolle in der Entwicklung und es geht nicht nur darum, etwas möglich schnell auf den Markt zu bringen. Dieser Prozess der Überprüfung und Begleitung ist dabei nicht erst durch das autonome Fahren entstanden, sondern hat sich in der Automobilentwicklung schon viele Jahre bewährt. Zudem handelt es sich beim autonomen Fahren um ein hochkomplexes technisches System, in das verschiedene Innovationen einfließen und an dessen Entwicklung **zahlreiche Personen** meist innerhalb eines Unternehmens sowie in **verschiedenen Unternehmen** beteiligt sind.

Dies steht im starken Gegensatz zur Entwicklung von **Deepfakes**. An ihrer Entwicklung (und an der Entwicklung von Deepfake-Detektionsmechanismen) sind zwar zunehmend auch große Tech-Unternehmen wie Google, Baidu, Facebook oder Microsoft beteiligt. Sie dominieren das

Feld jedoch (noch) nicht. Stattdessen sind – wie in vielen Feldern des maschinellen Lernens – **auch kleinere (private) Akteur*innen** bedeutend. Dazu gehören eigenständige Programmierer*innen, die ihren Code (häufig anonym) auf Open Source Code-Repositories wie GitHub und sozialen Netzwerken wie Reddit teilen. Sie sind oft nicht kommerziell motiviert, sondern durch eigene Neugier und einen ‚Open Source-Ethos‘. Daneben treiben auch professionellen Programmierer*innen in kleinen IT-Unternehmen, Start-ups oder an Forschungsinstituten die Deepfake-(Detektions-)Entwicklung voran. Dadurch ist die Entwicklung von Deepfakes deutlich dezentraler und liegt nicht hauptsächlich bei mächtigen, institutionellen Akteur*innen. Durch die fehlende institutionelle oder zentralistische Entwicklungsstruktur weist die Entwicklung von Deepfakes bislang **kaum Pfadabhängigkeiten** auf. Wichtige Kriterien für die Erstellung von qualitativ hochwertigen und realistischen Deepfakes sind **Geschwindigkeit und Kreativität**. Im Vergleich zum autonomen Fahren spielt der Aspekt von Sicherheit – sowohl im Sinne von Sicherheit von Leib und Leben als auch von Rechtssicherheit oder Sicherheit der Daten – in der Entwicklung keine größere Rolle.

a) Relevanz des Entwicklungskontexts für Governance

Diese Unterschiede im Entwicklungskontext des autonomen Fahrens und der Deepfakes haben entscheidenden **Einfluss auf die Wirksamkeit verschiedener Governancemechanismen**. Die Steuerung der Entwicklung von Deepfakes im Vergleich zum autonomen Fahren ist weniger geradlinig. **Aufgrund des informellen Charakters der Entwicklungsumgebung von Deepfakes ist es sehr schwierig, den Entwicklungsprozess von Deepfakes überhaupt zu steuern**, um sozial und ethisch wünschenswerte Innovationen zu erreichen. Eine Ethik-Kommission wie im Bereich des autonomen Fahrens hätte in diesem Fall keine Wirkung auf die Entwicklung. Auch technische Normen und entsprechende Prüforganisationen haben durch die Art der Entwicklung kaum Zugang, da viele Deepfakes im privaten Kontext erstellt werden. Sie könnten nur bei kommerziellen Deepfakes sowie bei in ‚formalen‘ Kontexten entwickelten Detektionssystemen wirken. Gleichzeitig besteht in diesen Kontexten aber die Gefahr, dass dadurch Innovation und Dynamik unverhältnismäßig gebremst werden.

Da es schwierig ist, die Entwicklung von Deepfakes zu beeinflussen, konzentrieren sich die meisten **Governancemechanismen auf deren Anwendung und Verbreitung** in unterschiedlichen Kontexten. Dabei ist es wichtig, Schlüsselakteur*innen und Infrastrukturen zu identifizieren (im Falle von Deepfakes insbesondere [die Betreiber*innen] soziale[r] Medien sowie Messenger-Dienste).

Beim **autonomen Fahren** dagegen sind die **Schlüsselakteur*innen** (zumindest in Deutschland) **bekannt** und bereits in die Governance eingebunden, unter anderem in verschiedenen Gremien wie dem Runden Tisch Automatisiertes Fahren, bei dem neben Behördenvertreter*innen auch Akteur*innen aus der Industrie und andere beteiligt sind. Die vorhandenen Governancemechanismen sind in diesem Feld schon lang etabliert, wie beispielsweise die Entwicklung nach Normen; es besteht eine starke **Pfadabhängigkeit**. Aufgrund dieser Ausgangslage ist es möglich und auch wirkungsvoll die **Governance auf die Entwicklung zu fokussieren**, beispielsweise durch technische Normen, Institutionen wie der TÜV und eine Forschungsförderung, welche normative und empirische Wissenschaften in die Technik-Forschung integriert.

Dieser Vergleich zeigt:

Der jeweilige **Kontext der Technologieentwicklung ist entscheidend**, um Governancemechanismen zu bestimmen, die ethische Überlegungen wirksam in Innovationsprozesse einbeziehen können – ob in der Entwicklung oder Anwendung. Anders gesagt: Die jeweiligen Entwicklungskontexte von Technologien als (Ansammlung) passiver Governancemechanismen beeinflussen, inwiefern Instrumente zur Regulierung der *Entwicklung* oder der *Anwendung* und *Verbreitung* einer Technologie wirksam sein können. Ein wichtiger Ansatzpunkt für Governance ist es daher, eine kontextsensible und interdisziplinär informierte Regulierung zu fördern, um herauszufinden, ob eine Konzentration auf den Entwicklungs- oder Anwendungskontext für die Regulierung vielversprechender ist.

Dabei sind folgende **Fragestellungen** mit jeweiligen Unterfragen für den Einsatz von Governancemechanismen zentral:

1. Wer sind die Schlüsselakteur*innen?

Welche Akteur*innen sind im Innovationsfeld vorhanden und wie mächtig sind sie jeweils? Welche Akteur*innen werden dadurch marginalisiert? Welche Innovationsprozesse werden dadurch gehemmt oder gar verdrängt und welche Interessen werden nicht berücksichtigt? Durch wen wird die Entwicklung vorangetrieben oder dominiert? Durch wen (Akteur*innen oder Plattformen) und wie wird die Technologie verbreitet?

2. Wer sind die Entwickler*innen und was motiviert sie?

Welche Interessen verfolgen die Entwickler*innen? Kann man die Entwickler*innen gut an den Verhandlungstisch holen oder bleiben sie eher heterogen und diffus?

3. Welche passiven Governancemechanismen bestehen?

Inwieweit gibt es bereits (technische Infra-)Strukturen und Institutionen sowie soziale Kontexte, welche die Wirkung verschiedener Governancemechanismen beeinflussen? Wer kontrolliert die vorhandenen Strukturen und in wessen Hand liegt die notwendige Infrastruktur? Welche Effekte haben sie auf die Technologieentwicklung? Können Pfadabhängigkeiten identifiziert werden und

wenn ja, welche Bedeutung haben diese für die Governance? Werden daraus bestehende Strukturen sichtbar, die genutzt werden können, um die Wirkung von neuen Governancemechanismen zu verstärken? Oder stehen die bisherigen Strukturen der Regulierung im Weg und wie können sie dann aufgebrochen werden? Wie verändert sich die bestehende Infrastruktur durch die Innovation und welche Folgewirkungen hat dies auf andere Bereiche?

4. Welche Interessen und Machtstrukturen sind vorhanden?

Sowohl in Technologien mit einer starken Pfadabhängigkeit wie auch in ‚neuen‘ Technologien, deren Entwicklungskontext in neuen und/oder diversen Feldern entsteht, sind vielfältige Interessen und Machtstrukturen vorhanden. Wer verfolgt welche Interessen? Durch welche Machtstrukturen wird das System getragen? Wer wird marginalisiert?

Eine genaue Analyse des Entwicklungskontexts einer technologischen Innovation erlaubt daher Rückschlüsse auf die voraussichtliche Wirksamkeit unterschiedlicher Governancemechanismen.

Bei Innovationen, deren Entwicklungskontext dezentral und informell ist, ist eine Konzentration auf die Regulierung und Beeinflussung der *Verbreitung* und des *Konsums* dieser Technologie vielversprechend.

Hierbei steht die Frage im Fokus, wer diese Technologie nutzt, mit welchem Motiv und über welche Infrastruktur die Menschen darauf zugreifen. Im Falle von Deepfakes wäre hier etwa eine stärkere gesetzliche Regulierung und Selbstregulierung von Plattformen und Messengern denkbar. Hierbei muss allerdings immer abgewogen werden, ob solch eine Regulierung die negativen Implikationen einer Technologie wirklich eingrenzen kann, oder inwieweit dies in andere Werte, wie beispielsweise die Meinungsfreiheit eingreift. Eine weniger invasive Herangehensweise wäre demnach, das Bewusstsein für technologische Risiken in der Öffentlichkeit zu erhöhen, denn eine systematische Medienkompetenzbildung für die Öffentlichkeit kann präventiv die Gefahren von Deepfakes reduzieren. Im Falle von Deepfakes kann dies sehr wirksam sein, bietet aber keinen vollständigen Schutz. Zudem muss ein **besonderes Augenmerk auf den Schutz von vulnerablen Gruppen** gelegt werden. Diese unterscheiden sich dabei je nach Anwendungskontext. So gehören im Fall von Deepfakes für Satire und Ironie Menschen mit eingeschränkten kognitiven Fähigkeiten oder geringen Sprachkenntnissen zu den vulnerablen Gruppen, da für sie unter Umständen die Fälschung nicht so gut erkennbar wird. Bei Deepfakes im pornografischen Kontext müssen

Governancebemühungen dagegen insbesondere Frauen und Minderjährige im Blick haben, die meist Opfer solcher Deepfakes werden.

Bei Technologien, die zentralisiert und industriell entwickelt werden, ist dagegen eine **Konzentration auf den *Entwicklungsprozess*** der Technologie vielversprechend.

So sind entsprechende Reflexionsprozesse im autonomen Fahren bereits in die Entwicklung eingebunden, beispielsweise durch die Einsetzung einer Ethik-Kommission, aber auch durch den Einfluss von Normen. Hierbei wird sowohl empirisches als auch normatives Wissen bereits in der Governance der Entwicklung berücksichtigt; zugleich wird die Integration normativer und empirischer Wissenschaften in die Technikforschung beispielsweise durch die Forschungsförderung vorangetrieben. Unsere Analyse hat zwar gezeigt, dass die Governance diesbezüglich noch ausbaufähig ist, aber auch, dass die Grundlage für eine solche Governance grundsätzlich im autonomen Fahren vorhanden ist und funktioniert. Eine solche Governance des Entwicklungsprozesses ist dabei potenziell auch auf andere Felder übertragbar, in denen dieser formalisiert und industriell geprägt ist.

Für die Governance der Entwicklung sollten normative und empirische Wissenschaften stärker in die Technikforschung integriert sowie bei der Governance berücksichtigt werden.

b) Berücksichtigung aller möglichen Regulierungskontexte

Auch wenn eine Konzentration entweder auf die Entwicklung oder die Verbreitung und den Konsum einer Technologie besonders vielversprechend ist, sollten für eine möglichst wirksame Governance **alle möglichen Regulierungskontexte berücksichtigt** werden. Es sollten auch solche nicht vernachlässigt werden, die aufgrund des Entwicklungskontexts weniger naheliegend erscheinen. So sollte das Potenzial von Mechanismen, die auf die Governance des *Entwicklungsprozesses* abzielen, auch bei **technologischen Innovationen** nicht außer Acht gelassen werden, **deren Entwicklungskontext dezentral und informell ist**.

Bei Deepfakes können beispielsweise eine stärkere Moderation oder sogar gesetzliche Regulierung von Open Source-Plattformen wie Github und Deepfake-Foren zentrale Governancemechanismen sein. Dies zielt stärker auf die Infrastruktur der Entwickler*innen ab als auf die Entwickler*innen an sich. Darüber hinaus könnte die Anregung zu einer ethischen Reflexion von Entwickler*innen in Bezug auf das Missbrauchspotenzial der von ihnen entwickelten Technologien und ihre Verantwortung die schädigende Wirkung von Deepfakes minimieren.

Dagegen sollten bei technischen Innovationen, die eher zentralisiert und formell entwickelt werden, auch Governancemechanismen berücksichtigt werden, die deren Verbreitung und Konsum beeinflussen. Zwar können hier Governancemechanismen, die auf die Technologieentwicklung abzielen, sehr effektiv sein. Dennoch muss **auch hier auf die tatsächliche Anwendung** in unterschiedlichen Kontexten geachtet und diese sollte je nach Kontext reguliert werden.

So ist beim autonomen Fahren auch die Gesetzgebung entscheidend, welche die Verwendung der Autopiloten im Straßenverkehr und damit die Anwendung und Verbreitung der Technologie des autonomen Fahrens reguliert. Anpassungen in der Straßenverkehrsordnung legen dabei fest, wie autonomes Fahren im Straßenverkehr aktuell realisierbar ist – und wie nicht. Darüber hinaus wirkt sich beispielsweise die rechtliche Klärung der Haftungsfragen im Falle eines Unfalls auch auf die Anwendung der Technologie aus: Es kann Nutzer*innen dazu veranlassen, entsprechende Funktionen bedachter zu verwenden, statt sie unzureichend zu überwachen – was in der Vergangenheit schon zu Unfällen geführt hat. Dementsprechend spielt auch die Governance der Anwendungskontexte eine Rolle für die Governance des autonomen Fahrens.

c) Chancen nutzen und bestehende Strukturen hinterfragen

Dadurch, dass die Technologie von Deepfakes im Vergleich zum autonomen Fahren noch recht jung ist und ihre Entwicklungsumgebung gerade erst aufgebaut wird, sind Strukturen bislang weniger festgefahren. Somit könnte die Chance bestehen, dass **größere Spielräume vorhanden sind, durch Governance ethisch Implikationen zu berücksichtigen und die Technologie sozial wünschenswert** zu gestalten. Solche Chancen sollten aktiv genutzt werden.

Im Falle von Deepfakes und anderen Online-Desinformationsphänomenen geschieht das etwa in der aktuellen politischen und rechtlichen Debatte um Plattformregulierung. Bei diesen Bemühungen betreten nationale und supranationale Gesetzgebungsorgane meist Neuland, und es gibt Spielräume für neue Mechanismen etwa der begleiteten Selbstregulierung (wie der ‚EU Code of Practice Against Desinformation‘), aber auch für deren kritische Evaluation und die Etablierung neuer Governancemechanismen wie dem EU Digital Services Act. Auch innovative Governancemechanismen wie etwa Wettbewerbe, Challenges oder Hackathons, die unter anderem von großen Tech-Unternehmen ausgerufen werden, können hier maßgeblich zur Entwicklung beitragen. Da viele Deepfakes von Privatpersonen erstellt werden, besteht zudem Potenzial, neue Strukturen zwischen einzelnen, privaten Akteur*innen zu schaffen und diese ethisch und sozial wünschenswert auszugestalten. So könnten in Online-Communities und Foren wie beispielsweise Reddit oder GitHub Aufklärungskampagnen den Austausch zwischen den Nutzer*innen und eine ethische Folgenabschätzung fördern. Dies würde das Bewusstsein für individuelle Verantwortung und die Missbrauchsprävention stärken. Dadurch kann ein gesamtgesellschaftlicher Blick für Chancen und Risiken durch Deepfakes geschaffen werden, und es können neue Pfadabhängigkeiten gezielt ethisch und sozial wünschenswert geformt werden.

In anderen **Feldern, in denen das sozio-technische System sehr stabil** ist, darf das Vorhandensein beispielsweise einer starken Entwicklungstradition nicht dazu verleiten, diese **Ausgangslage** (passive Governancemechanismen wie Entwicklungstraditionen oder -infrastrukturen) als gegeben hinzunehmen. Stattdessen muss sie **im Hinblick auf ihre Zukunftsfähigkeit kritisch hinterfragt** werden: Wo können und müssen Strukturen aufgebrochen werden, um ethische und gesellschaftliche Aspekte zu berücksichtigen? Wie können neuen, nicht etablierten Entwickler*innen Zugänge zum Feld verschafft werden? Wie kann hier agiert werden, um so nicht nur die Innovation zu befördern, sondern auch ethische und gesellschaftliche Aspekte zu berücksichtigen?

Beispielhaft ist hier im Feld des autonomen Fahrens die große Rolle des Individualverkehrs zu nennen, welche durch die entsprechende Infrastruktur verfestigt ist. Dadurch wird auch die Entwicklung des autonomen Fahrens häufig in die Richtung des Individualverkehrs gedacht. Zwar wird auch an Shuttlekonzepten gearbeitet, doch sind diese Bemühungen durch die vorhandene Infrastruktur meist limitiert. Würden sich neue Konzepte weniger eng am Individualverkehr orientieren, könnten Nachhaltigkeitsziele und Fragen von Gerechtigkeit einfacher geklärt werden – durch effizientere Verkehrslösungen, die nicht an den individuellen Autobesitz gebunden sind. Darüber hinaus ist auch die Macht der Automobilhersteller ein großer Faktor im Feld des autonomen Fahrens, da sie als Akteur*in häufig in Governancemechanismen eingebunden ist. Dies gibt den Automobilherstellern die Möglichkeit, ihre Interessen im Entwicklungsprozess prominent zu äußern und durchzusetzen, die primär ökonomischer Natur sind und nicht vorwiegend von ethischen und gesellschaftlichen Abwägungen geleitet werden. Durch ein Hinterfragen der Strukturen könnte dies aufgebrochen werden, wenn die Governance eben nicht nur auf bislang existierenden Strukturen basiert, sondern neue Wege geht und beispielsweise neue Akteur*innen einbezieht. Die Forschungsförderung zeigt bereits eine Möglichkeit, wenn dort beispielsweise spezifisch Shuttlekonzepte gefördert werden. Diese Bemühungen sollten noch weiter ausgebaut werden und vor allem interdisziplinär gedacht werden. Insgesamt braucht es Raum für kreative Ideen, um Innovationsprozesse zukunftsfähig zu gestalten.

d) Mündige Technologienutzung durch Aufklärung und Bildung

Die Kontrastierung der beiden Felder des autonomen Fahrens und der Deepfakes bieten Ansatzpunkte für eine effektive Innovationsgovernance in unterschiedlichen Kontexten. In einem Punkt sehen wir jedoch auch eine große Übereinstimmung zwischen beiden Feldern:

Es besteht **großes Potenzial in einer breiten Aufklärung** und in der Anregung von Reflexionsprozessen zu den Herausforderungen und Konsequenzen des technologischen Wandels sowie über die ethischen und gesellschaftlichen Implikationen technologischer Innovationen. Sowohl für Deepfakes als auch im Feld des autonomen Fahrens hat sich gezeigt, dass **ein besseres Verständnis der Chancen und Risiken der Innovationen einen positiven Effekt auf die Entwicklung und Anwendung der Technologien** im Sinne einer stärkeren ethischen und gesellschaftlichen Wünschbarkeit hätte.

Zentral ist dabei:

1. Das Bewusstsein für den Einfluss von technologischem Wandel auf die Gesellschaft zu erhöhen.
2. Das Bewusstsein für die Ambiguität und Werthaltigkeit von Technik zu erhöhen.
3. Einen mündigen und kritischen Umgang mit Technologie fördern, der es ermöglicht bestimmte technische Entwicklungen bezüglich ihrer grundsätzlichen Sinnhaftigkeit und Wünschbarkeit zu hinterfragen.

Dies hat mehrere positive Effekte:

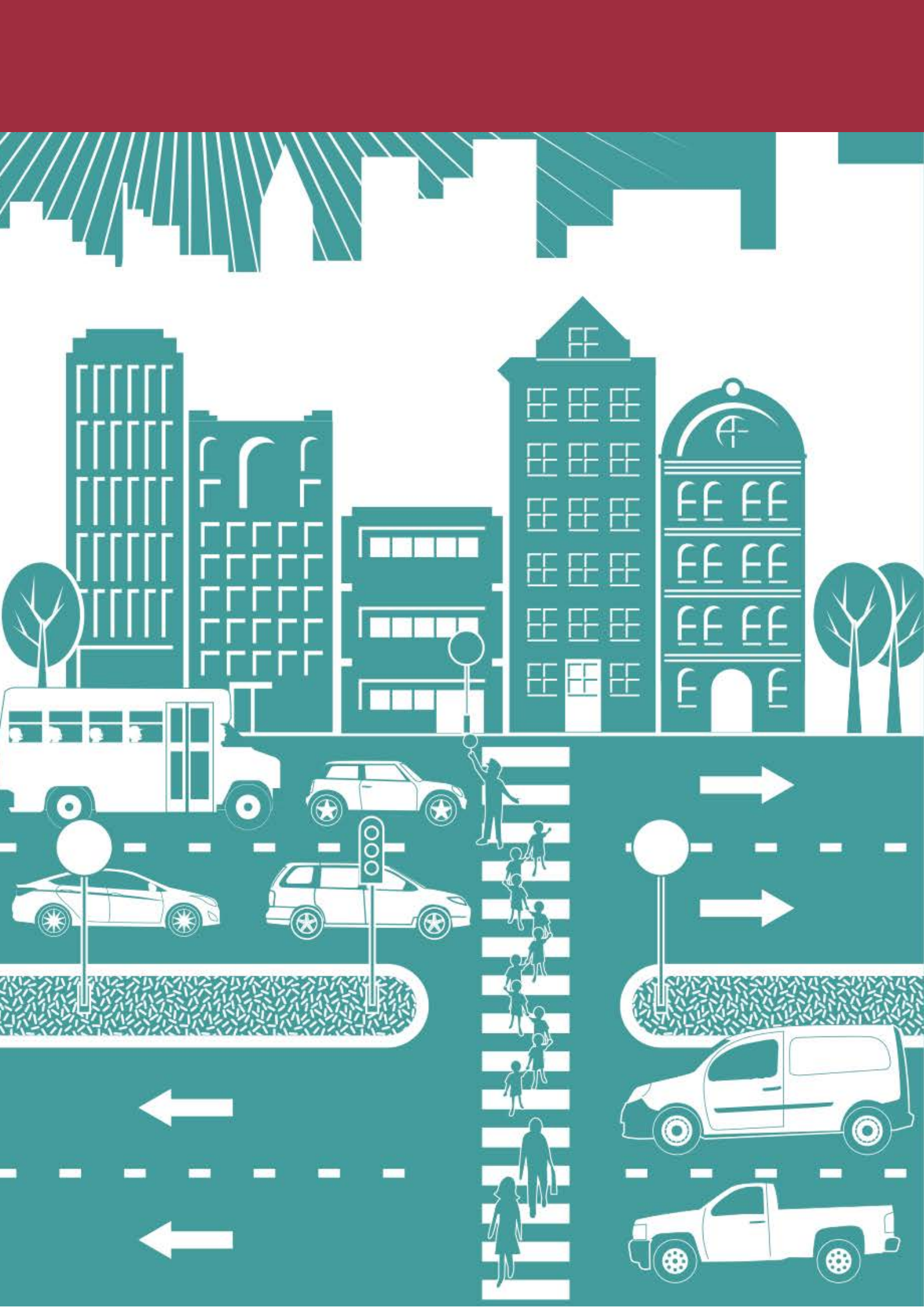
1. Es fördert einen aufgeklärten Diskurs, sodass die Debatte über Technologien wie dem autonomen Fahren nicht oberflächlich stattfindet. Eine tiefgreifendere Debatte und ein größeres Technikverständnis wirken **unbegründeter Skepsis und irrationalen Befürchtungen entgegen**, erhöhen aber auch das **Bewusstsein für die tatsächlichen Problematiken und Gefahren** einer Technologie.
2. Eine solche tiefgreifendere Debatte beeinflusst auch die Entwicklung von Technik und hätte damit wiederum einen positiven Effekt auf deren **Akzeptabilität (und damit deren ethische und gesellschaftliche Wünschbarkeit)**, da Entwickler*innen ethisch informiert und interdisziplinär ausgerichtet wären und die Technikentwicklung stärker gesellschaftlich rückgebunden werden könnte.
3. Ein solches Bewusstsein würde Anwender*innen/Nutzer*innen ermöglichen **verantwortlich(er) und autonom(er) mit neuen Technologien umzugehen** und hat dadurch das Potenzial, gewissen Risiken vorzubeugen.

Wünschenswert ist in diesem Zusammenhang eine **Zivilklausel**, wie sie an vielen Universitäten verbreitet ist, die eine Selbstverpflichtung der Institutionen darstellt, nur für zivile Zwecke zu forschen. Sie führt zu einem stärkeren Bewusstsein für die Dual-use Problematik und könnte dementsprechend auch für andere Institutionen wegweisend sein.

Eine Aufklärung der Öffentlichkeit für Chancen und Risiken von Technologien ist dabei ein **wenig invasiver** Governancemechanismus (etwa im Vergleich zu rechtlicher oder technischer Normierung). Gleichzeitig kann sie ein **starkes Instrument für eine ethisch und sozial wünschenswertere Technikentwicklung** sein – sowohl bei Technologien, deren Entwicklungsprozess durch Governance direkt beeinflusst werden können, als auch bei solchen, bei denen Entwicklung schwer durch Governance greifbar ist.

Um diese Stärkung der öffentlichen Debatte auch in die Entwicklung rückzubinden, müssen Möglichkeiten geschaffen werden, wie die Ergebnisse aus der öffentlichen Diskussion wieder in die Entwicklung einfließen können. Ziel dabei muss es dabei auch sein, die Wünschbarkeit bestimmter Entwicklungen generell (auch politisch) zu hinterfragen, und damit ethisch problematische Entwicklungen oder Entwicklungen mit hohem Schadenspotenzial entweder stark zu regulieren oder aber nicht zu fördern.

Diese Aufklärung kann und sollte dabei **nicht nur auf einer Ebene** stattfinden, sondern sollte unter anderem in der formalen und universitären Bildung erfolgen. Sowohl durch das Curriculum technischer als auch nicht-technischer Disziplinen sollte ein Bewusstsein für ethische Fragestellungen und Kenntnis über den Umfang und die Grenzen der eigenen Verantwortung geschaffen werden. Darüber hinaus sollte hier Interdisziplinarität sinnvoll eingebunden werden. Für eine breite Wirkung ist es darüber hinaus sinnvoll, auch berufliche Weiterbildungsangebote einzubeziehen, ebenso wie Bildungsangebote für eine breite, interessierte Öffentlichkeit.



III. Handlungsempfehlungen Autonomes Fahren

Ethische, rechtliche und gesellschaftliche Begleitung der Entwicklung

- Die Technik des autonomen Fahrens verspricht große Vorteile (Sicherheit, Nachhaltigkeit). Mit ihr gehen aber auch Nachteile und Risiken einher (Gerechtigkeit, Machtverteilung). Eine rein punktuelle ethische Bewertung reicht nicht aus. Mit dem Fortentwickeln der Technik verändern sich die ethischen Einschätzungen und neue normative Fragen entstehen. Daher ist es notwendig, die laufende Entwicklung ethisch und gesellschaftswissenschaftlich zu begleiten. Verschiedene Komponenten der Technik (z.B. Vernetzung) und Szenarien (z.B. Mischverkehr) müssen hinsichtlich ihrer Auswirkungen untersucht werden und die Chancen und Risiken gesellschaftlich ausgehandelt werden.
- Nicht nur Fragen der (gesellschaftlichen) Akzeptanz dürfen die Erforschung gesellschaftlicher, ethischer und rechtlicher Aspekte leiten, sondern genauso Fragen der (ethischen) Akzeptabilität.
- Wenn die Entwicklung weiter Richtung Stufe vier und fünf des autonomen Fahrens reift, werden hier neue ethische, rechtliche und gesellschaftliche Herausforderung entstehen. Diese müssen laufend untersucht werden, damit die Governance frühzeitig reagieren kann.

Ethik-Kommission Automatisiertes und Vernetztes Fahren

- Aufgrund der laufenden Entwicklung der Technik ist es notwendig, die Arbeit der Ethik-Kommission weiter zu führen, um auch neuen Herausforderungen begegnen zu können.
- Für eine ethische Beurteilung muss auch ethische Forschung vorhanden sein. Deswegen ist es entweder notwendig, die Ethik-Kommission selbst mit Forschungskapazität auszustatten, oder anderweitig ethische Forschung zu fördern, auf deren Ergebnisse sich die Ethik-Kommission in ihren Entscheidungen stützen kann.
- Ethische Empfehlungen sind häufig nicht 1:1 von Techniker*innen in die Entwicklung zu übertragen, sondern stellen eher einen Reflexionsmoment da. Dennoch muss ein Weg gefunden werden, wie die Ergebnisse in der Praxis stärker berücksichtigt werden können. Hierfür wird ein Dialog zwischen Forscher*innen aus den Geistes-, Rechts- und

Sozialwissenschaften und Techniker*innen und mehr wechselseitiges Verständnis für die jeweils andere Denkweise benötigt. Der Dialog kann durch interdisziplinäre Projekte gefördert werden, ein tiefergehendes wechselseitiges Verständnis durch interdisziplinäre Studiengänge, die frühzeitig bereits beide Wissensarten vernetzen (siehe auch Empfehlung zu Bildung).

- Derzeit wird nicht nach außen kommuniziert, dass die Ergebnisse der Ethik-Kommission kein genuin ethisches Urteil sind. Stattdessen ist der Bericht schon das Ergebnis einer gesellschaftspolitischen Aushandlung. Alternativ kann das Gremium als reine Ethik-Kommission eingesetzt werden, die nur mit wissenschaftlichen Akteur*innen besetzt ist und weniger mit Interessensvertreter*innen verschiedener Gruppen wie der Automobilhersteller.
- Der Einfluss von Diskursen sollte stärker berücksichtigt werden, da sonst entscheidendes Potenzial verloren geht. So hat die Ethik-Kommission gezeigt, wie ein zu starker Fokus auf der *Akzeptanz* des autonomen Fahrens durch die Öffentlichkeit die Bandbreite der Diskussion ethischer Aspekte einschränkte. In Zukunft sollte der Schwerpunkt daher auf *Akzeptabilität* gelegt werden.
- Die Besetzung einer Ethik-Kommission sollte diverser erfolgen, sodass die Vielfalt der Gesellschaft abgebildet wird. Unter anderem sollten mehr Frauen beteiligt werden. Zudem müssen insgesamt mehr Ethiker*innen berufen werden, wenn ein dediziert ethisches Urteil ausgehandelt werden soll.
- Aktuell sind die Ergebnisse der Ethik-Kommission nicht bindend. In Zukunft sollte durch die Politik stärker definiert werden, ob bestimmte Ergebnisse für welche Adressaten bindend sein sollten bzw. ihren Weg in Gesetzesvorlagen finden sollten.

Gesetze

- Auch wenn Gesetze kaum die Frühphase der Entwicklung betreffen, ist es dennoch wichtig, durch Gesetzgebung Rechtssicherheit zu schaffen. Hierdurch kann sichergestellt werden, dass Technik, die in Deutschland verwendet wird (einschließlich importierter Technik), den allgemeinen ethischen, rechtlichen und gesellschaftlichen Maßgaben entspricht.
- Bei der Gesetzgebung sollten die Erkenntnisse rechtlicher, ethischer und gesellschaftswissenschaftlicher Forschung zum autonomen Fahren kontinuierlich einbezogen werden.

Forschungsförderung

- Eine stärkere Integration normativer (insbesondere ethischer und rechtlicher) und gesellschafts- und sozialwissenschaftlicher Forschung in die Entwicklung des autonomen Fahrens ist notwendig. Es muss ein Fokus auf die komplexen ethischen Abwägungen beispielsweise zwischen Sicherheit, Transparenz und Datenschutz gelegt werden. Dazu ist eine entsprechende Forschungsförderung nötig. Gerade dort, wo die Ambivalenzen einer technischen Innovation hervortreten, müssen neue Entwicklungsschritte stets im Hinblick auf Wertkonflikte und Abwägungsprozesse bearbeitet werden. Dies ist eine genuine Aufgabe ethischer Forschung. Ebenso wichtig ist der Einbezug rechtlicher und empirischer Forschung über die gesellschaftlichen Auswirkungen autonomen Fahrens.
- Normative und empirische Wissenschaften müssen als Governance-Mechanismen anerkannt werden, da die Forschung selbst Wirkung auf das Feld und in die Gesellschaft hat. Das kann durch die Forschungsförderung der Politik strategisch mehr genutzt werden
- Die nationale, aber auch supranationale Forschungsförderung sollte eine interdisziplinäre Verzahnung auch in der Praxis realisieren. Förderlinien, die ethisch und gesellschaftlich höchst relevante Themen ansprechen, müssen dies auch in der Ausschreibung berücksichtigen. In den Vergabeverfahren sollten zudem auch tatsächlich Projekte zum Zuge kommen, die ethische, rechtliche und gesellschaftliche Aspekte (interdisziplinär) bearbeiten.
- Es gibt noch zahlreiche Projekte mit hoher ethischer, gesellschaftlicher und rechtlicher Relevanz, die keine eigene Forschung aus diesen Disziplinen einbinden. Diese Lücken sollten kompetent aufgedeckt und geschlossen werden.
- Das bisherige Potenzial der Forschungsförderung als Governancemechanismus wird noch nicht ausgeschöpft. Die Forschungsförderung kann stärker als Governancemechanismus genutzt werden, um ethische und gesellschaftliche Fragen einzubeziehen, sowie um bestehende Strukturen zu hinterfragen. So können Alternativen zu bestehenden Forschungspfaden gefunden und starre Entwicklungstradition aufgebrochen werden. Gerade in Hinblick auf die Herausforderungen des 21. Jahrhunderts mit Zielen wie Nachhaltigkeit ist dies ein zentrales Anliegen.

Normung

- Der Entwicklungskontext des autonomen Fahrens zeigt, dass technische Normen ein wichtiger Bestandteil der Governance der Entwicklung von Fahrzeugen sind. In der Praxis ist es jedoch bisher schwierig, ethische Werte und Aspekte in technische Normen zu

übersetzen. Um Normung als Governancemechanismus zu erschließen, ist weitere interdisziplinäre Forschung nötig. Diese ist in der Lage herauszuarbeiten, wie und zu welchem Grad ethische Aspekte in technische Normen übersetzt werden können, sodass diese sowohl den ethischen und gesellschaftlichen Ansprüchen gerecht werden als auch für die Entwickler*innen anwendbar sind.

- Die Werte, die in bisherige Normen eingeflossen sind, sollten stärker hinterfragt werden. So stellt sich beispielsweise die Frage, für wen die Sicherheit gilt, die aktuell durch Normen gewährleistet werden soll.
- In Normen sollte zudem ein breiteres Spektrum von ethischen Werten einbezogen werden. Ein einseitiger Fokus auf Sicherheit ist nie zielführend, da Sicherheit immer in Konkurrenz zu anderen Werten steht. Die konkrete Abwägung von Werten in unterschiedlichen Kontexten macht intensive Forschungsbemühungen nötig. Zudem müssen diversere Akteur*innengruppen, wie Ethiker*innen, aber auch Verbraucher*innenorganisationen und Nutzer*innen in die Entwicklung der Normen mit einbezogen werden.
- Eine Forschungslücke ist bislang, wie eine (noch stärkere) Berücksichtigung ethischer und gesellschaftlicher Implikationen in der Begleitung der Entwicklung durch den TÜV aussehen kann. Hier ist auch der TÜV in der Verantwortung, seine Rolle zu reflektieren und ethische und gesellschaftliche Fragen stärker in seine Prozesse mit einzubeziehen.
- Aufgrund der bisherigen Schwierigkeiten, ethische Werte in Normen einzubeziehen, darf nicht ausschließlich auf Governance durch Normung gesetzt werden. Stattdessen braucht es flankierende Maßnahmen wie etwa eine allgemeine Stärkung ethischer Kompetenzen insbesondere bei Entwickler*innen.

Entwicklungstradition

- Es ist sinnvoll und nötig, die Entwicklungstradition im autonomen Fahren dort zu hinterfragen und aufzubrechen, wo sie nachhaltige und gesellschaftlich wünschenswerte Entwicklung blockiert. Dazu wird ein besseres Verständnis über Pfadabhängigkeiten in der Entwicklung und deren Auswirkung auf den Einsatz und die Möglichkeiten von Governance benötigt. Dies baut auf entsprechender Forschung auf.
- Insbesondere der alleinige Fokus auf den Individualverkehr muss erweitert werden, da dieser eine nachhaltige und gerechte Umsetzung autonomen Fahrens erschwert. Gerade der Fokus auf Sharing-Konzepte und eine neue Art des Massenverkehrs sollte weiter gestärkt werden. Da hieran große wirtschaftliche und gesellschaftliche Veränderungen hängen, braucht dieser Wandel die Basis interdisziplinärer Forschung, die mögliche Folgen evaluiert und ethische, rechtliche und gesellschaftliche Fragen frühzeitig erörtert.

Hierbei spielt insbesondere die Forschungsförderung eine zentrale Rolle. Des Weiteren sollten Verbindungen zu wirtschaftlichen Partner*innen hergestellt und sowohl Start-Ups aber auch etablierte Hersteller*innen mit eingebunden werden. Gleichzeitig werden neue Konzepte für die Innenstädte benötigt, die auf Basis ethischer und gesellschaftlicher Ansprüche erfolgen müssen.

- Dort wo sie sinnvoll ist, sollte die Entwicklungstradition des autonomen Fahrens aber auch als Ressource genutzt werden, um wirksame Governance zu ermöglichen. So ermöglicht die Normung einen guten Eingriff in die Entwicklung der Fahrzeuge.

Diskurse

- Diskurse müssen als Governancemechanismus anerkannt werden. Sie beeinflussen die Entwicklung zentral mit, da durch die öffentliche Meinung der Einsatz einer Technik in eine bestimmte Richtung gelenkt werden kann oder sogar ein bestimmter Aspekt komplett in Frage gestellt und so verhindert werden kann.
- Governancemechanismen sollten nicht nur nach dem öffentlichen Stimmungsbild ausgerichtet werden. Stattdessen muss hier genauer zwischen Akzeptanz und Akzeptabilität der Technik abgewogen werden. Es kommt hier auf eine fundierte Information der Bürger*innen über Chancen und Risiken der Technologie an. Hierbei ist nicht nur die Politik in der Pflicht, sondern auch andere Akteur*innen bspw. aus der Wirtschaft und Wissenschaft.
- Darüber hinaus sollten Wege gefunden werden, wie Betroffene und Nutzer*innen sinnvoll in die Governancemechanismen mit eingebunden werden können.

Partizipation

- Partizipation der Bürger*innen an der Entwicklung kann und muss ein Mittel sein einen tatsächlichen Einbezug von Nutzer*innen und Betroffenen herzustellen.
- Fragen von Partizipation werden zukünftig zu neuen und neu zu gestaltenden Governancemechanismen werden. Für das Feld des autonomen Fahrens ist hier – im Sinne der Verzahnung von Governancemechanismen – auch die Forschungsförderung in der Pflicht, die Erforschung und Entwicklung guter, gerechter und sinnvoller Formate der Partizipation zu fördern. Gleichzeitig ist es nötig, dass durch die Politik Partizipationsmöglichkeiten geschaffen und auch Regulierung auf den Weg gebracht werden, wie diese in die wirtschaftlichen Prozesse eingebunden werden kann.

Infrastruktur

- Im Fall des autonomen Fahrens entfaltet das gesamte soziotechnische System der Automobilität als passiver Governancemechanismus eine Wirkung auf die zukünftige Technologieentwicklung. Dies muss bei der Entscheidung berücksichtigt werden, welche Governancemechanismen zur Steuerung und Regulierung eingesetzt werden.
- Infrastruktur ist prägend für die neue Technik und hat Machtimplikationen – wer die Kontrolle über die Infrastruktur hat, kontrolliert die Mobilität großer Gruppen, hat Zugriff auf große Datenmengen und somit Wissen über das Verhalten vieler Menschen. Damit einher gehen auch Marktzugänge und somit ökonomisches Potenzial. Es ist daher wichtig, die Rolle von Infrastruktur stärker in den Fokus zu rücken. Sie darf nicht nur als gegeben hingenommen werden, sondern sollte auch hinterfragt und gegebenenfalls aufgebrochen werden. Wie kann die bestehende Infrastruktur sinnvoll eingebunden werden? Wo muss sie verändert werden oder neue Infrastruktur geschaffen werden, sodass eine möglichst gleichberechtigte Teilhabe an Mobilität in der Zukunft gewährleistet werden kann? Dies betrifft die Dateninfrastruktur des vernetzten Fahrens ebenso wie die Straßen selbst, und hier insbesondere die Frage, ob auf diesen nur autonome Fahrzeuge fahren dürfen oder auch weiterhin von Menschen gesteuerte Fahrzeuge.

Runder Tisch Automatisiertes Fahren

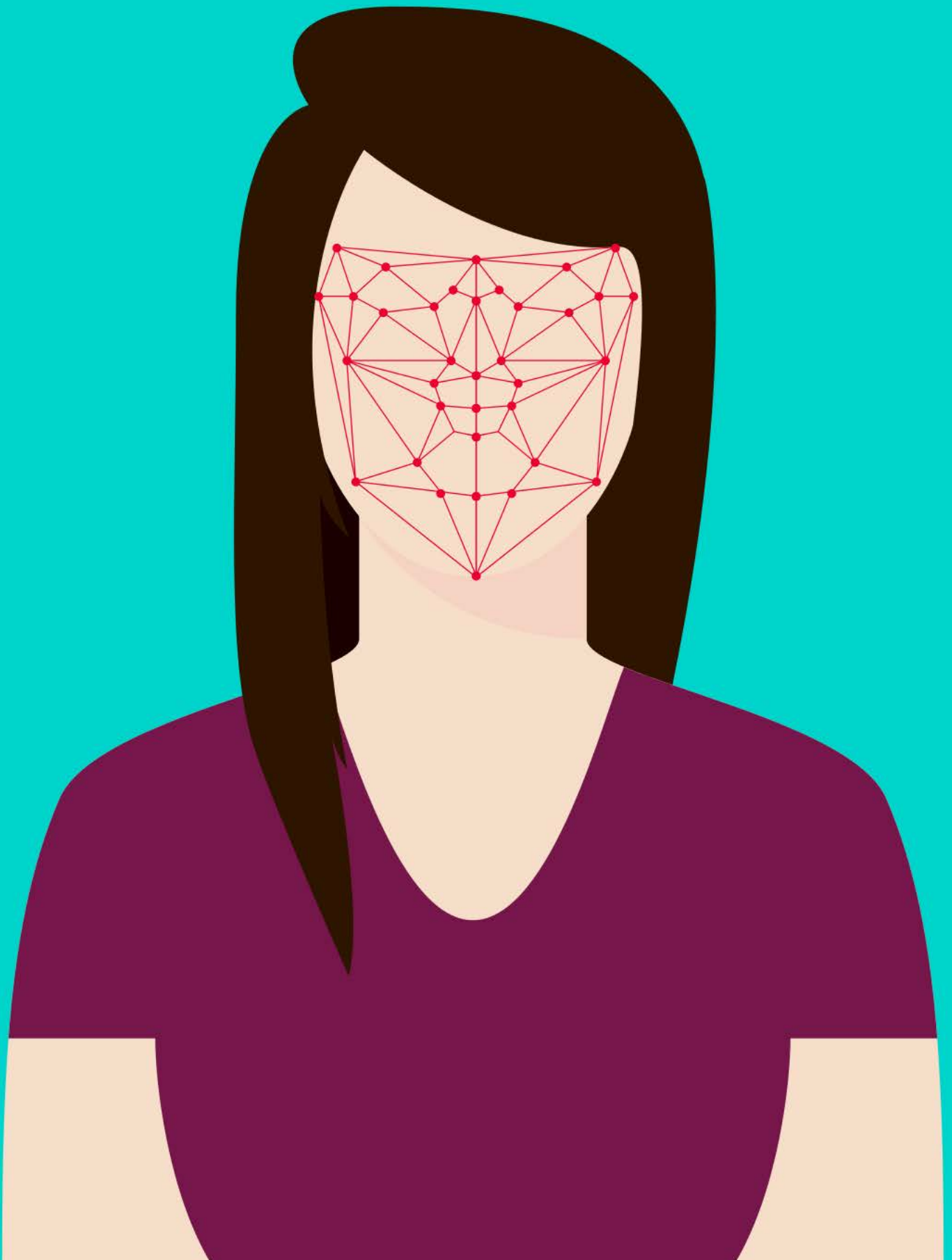
- Durch die zentrale Vernetzung wichtiger Akteur*innen im Feld des autonomen Fahrens hat dieses Gremium das Potenzial, wichtige Aspekte des autonomen Fahrens zu diskutieren. Durch die breite Beteiligung aus Wirtschaft, Forschung, Regulierung usw. sitzen hier zentrale Akteur*innen gemeinsam an einem Tisch und können sich und können sich wechselseitig gut verständigen. Jedoch sollten ethische Aspekte stärker berücksichtigt werden als bisher.
- Zudem sollte mehr Transparenz über die Arbeit und den Einfluss des Gremiums und die Rolle der beteiligten Akteur*innen geschaffen werden.

Bildung

- Die Angst vor vermeintlich schwerwiegenden ethischen Problemen wie beispielsweise den häufig diskutierten Dilemma-Situationen lässt andere ethisch relevante Aspekte oft in den Hintergrund rücken. Dahingehend ist eine öffentliche Aufklärung nötig, sodass es für

Nutzer*innen und Betroffene möglich ist, zwischen tatsächlichen Chancen und Risiken und überzogenen Erwartungen und Schwarzmalerei zu unterscheiden.

- Durch Studiengänge, die technische mit sozialwissenschaftlichen, ethischen und rechtlichen Fächern verbinden, können wechselseitiges Verständnis und Respekt als Basis für einen interdisziplinären Austausch geschaffen werden. Darüber hinaus sollten ethische Aspekte in Studiengänge wie Informatik, Künstliche Intelligenz, Ingenieurwissenschaften, Fahrzeugtechnik usw. integriert werden.
- Eine Stärkung ethischer Kompetenzen insbesondere bei Entwickler*innen könnte zum Beispiel die Berücksichtigung ethischer Aspekte in Normungsgremien fördern oder sogar zu einem Hinterfragen bestehender Entwicklungstraditionen und des bestehenden soziotechnischen Systems führen.
- Gleichzeitig würde ein besseres technisches Verständnis der Sozial- und Geisteswissenschaftler*innen sowie Ethiker*innen dazu führen, dass Entwickler*innen besser adressiert werden könnten. Gleichzeitig könnten die technischen Prozesse besser verstanden und erforscht werden.



IV. Handlungsempfehlungen Deepfakes

Kein generelles Verbot

- Ein komplettes und allgemeines Verbot von Deepfakes ist nicht sinnvoll. Es kann auch gefährlich sein, da es auch legitime oder sogar demokratiefördernde Nutzungen von Deepfakes gibt. Ein Verbot würde zudem die Meinungsfreiheit unzulässig einschränken, denn viele Deepfakes schaden nicht per se der Öffentlichkeit, sondern tragen auch zur öffentlichen Debatte bei. Dies umfasst vor allem Deepfakes in den Anwendungskontexten von Satire und Parodie, Kunst, Bildung, Aktivismus, und Assistenztechnologie. Darüber hinaus wäre ein umfassendes Verbot aufgrund der Vielzahl und verteilten Natur von Deepfake-Entwickler*innen und ihren Verteilungskontexten auch nicht umsetzbar.
- Es sollten jedoch alle Deepfakes in den genannten Anwendungsbereichen (bis auf im Bereich der Strafverfolgung von Kinderpornografie) als solche markiert werden, um Täuschungsabsichten entgegenzuwirken und Transparenz für die Betrachtenden herzustellen.
- Darüber hinaus muss es für jeden Anwendungskontext spezifische Regulierungsdiskussionen geben, die eine ethische und gesellschaftliche Folgenabschätzung voraussetzen.

Entwicklungs-Infrastrukturen und sozialer Kontext

- Diskurse und Wertvorstellungen zu ändern, ist ein langer und schwieriger Prozess und wird im Fall von Deepfakes durch die Anonymität und fehlende Institutionalisierung vieler Entwickler*innen weiter erschwert. Nichtsdestotrotz sollte versucht werden, das Bewusstsein der Entwickler*innen für die ethischen Folgen und die Werthaltigkeit ihrer Innovationen zu schärfen. Insbesondere die Erkenntnisse normativer Wissenschaften, also Ethik und Recht, sollten stärker in die Arbeit von Entwickler*innen auch auf Open Source-Plattformen einfließen. Dies könnte über Programme der Erwachsenenbildung erfolgen sowie direkter über Aufklärungsinitiativen und Peer-to-peer Moderation auf den Plattformen selbst.
- Förderlich wäre es, die Rolle von Moderator*innen auf Plattformen wie GitHub und Reddit zu stärken. Durch die Ausweitung von Regulierungsmaßnahmen könnten Plattformen stärker für rechtswidrige Inhalte zur Verantwortung gezogen werden.

- Förderprogramme könnten den weiblichen Nachwuchs in der Entwicklung stärken und somit zu mehr Geschlechtergerechtigkeit im Innovationsbereich beitragen.
- Der Open Source-Gedanke ist grundsätzlich ein wichtiger Governancemechanismus, der eine ethisch wünschenswerte Technologieentwicklung befördern kann. Jedoch muss der Diskurs um Open Source in Bezug speziell auf synthetische Medien – unter Berücksichtigung der Wissenschaftsfreiheit – kritisch überprüft werden, da es sich um ‚verbotenes Wissen‘ und ‚dual use‘-Anwendungen handeln könnte.
- Eine stärkere Regulierung von Suchmaschinen kann wirksam sein, damit schädliche Deepfake-Foren – insbesondere solche zur Erstellung nicht-einvernehmlicher Pornografie – nicht mehr über einfache Suchanfragen auffindbar sind.
- Eine Regulierung oder sogar ein Verbot von Foren wie 4chan oder 8chan, auf denen Deepfakes erstellt und verbreitet werden, kann der Verbreitung von Desinformation entgegenwirken und würde nicht im Widerspruch zu Meinungsfreiheit stehen, wenn diese (häufig sogar explizit) damit werben, sich nicht an (inter-)nationale Gesetze zu halten. Allerdings führt dies meist dazu, dass Nutzer*innen und Inhalte auf andere Plattformen migrieren. Daher wären eine Sperrung und ein Verbot solcher demokratiefeindlichen Plattformen vermutlich nur vorübergehend effektiv.

Infrastruktureller und sozialer Kontext der Verbreitung

- Es muss hinterfragt werden, wann die Anonymität auf sozialen Medien legitim ist und wie mit anonymen Posts umgegangen werden sollte.
- Gesellschaftliche und individuelle Tendenzen wie Frauenhass und Verschwörungsideologien, die hinter der Verbreitung vieler Deepfakes stehen, sollten durch Bildung, Aufklärung und zivilgesellschaftliches Engagement bekämpft werden.
- Suchmaschinen sollten verpflichtet werden, ihnen gemeldete klar rechtswidrige Deepfakes und damit zusammenhängende Seiten als solche kenntlich zu machen und algorithmisch herunterzubewerten bzw. aus den Suchergebnissen zu löschen. Dies würde die Transparenz für Benutzer*innen erhöhen, Betroffene schützen und Deepfakes anders positionieren
- Inhalte auf Messengern sind aufgrund ihrer Privatheit bislang schwer zu regulieren. Durch die Änderung von technischen Spezifika kann die Verbreitung von Deepfakes reduziert werden, beispielsweise durch eine Einschränkung der Gruppengröße privater Chats oder einer Beschränkung der Weiterleitungsmöglichkeiten von Nachrichten.

Regulierungen auf nationaler, supranationaler und internationaler Ebene

- Nötig sind gezielte Gesetze und Regulierungen von schädlichen Deepfakes auf nationaler und supranationaler Ebene, die insbesondere soziale Medien und Suchmaschinenbetreiber betreffen, da gerade pornographische und politische Deepfakes insbesondere über diese Kanäle verbreitet werden. Auf diesen Plattformen sowie über Suchmaschinen sollten Deepfakes, die klar rechtswidrig sind, nicht auffindbar sein.
- Die Entscheidungshoheit von Plattformen über die Legalität von Inhalten muss eingeschränkt werden, um die Machtkonzentration von Plattformen zu verringern. Das deutsche NetzDG überträgt hier viel Verantwortung an private Unternehmen. Das aktuelle Verfahren birgt demokratieschädliches Potential, da der Staat wenig Einfluss auf die Entscheidungen von privaten Unternehmen hat.
- Die aktuellen nationalen Gesetze bilden zudem neue Formen von Desinformation, zu der auch Deepfakes gehören, noch nicht ausreichend ab.
- Die EU kann bei der Regulierung von Hass, Hetze und Desinformation im Netz eine Vorreiterrolle einnehmen. Die von der EU-Kommission im Digital Services Act (DSA) vorgesehene Einführung von ‚vertrauenswürdigen Markierer*innen‘ (‚Trusted Flaggers‘) illegaler Inhalte kann zu einer schnelleren Löschung beitragen.
- Für eine Löschung von Inhalten sollte eine Erklärung gegeben werden. Erhöhte Transparenz schärft die Entscheidungsgrundlage, auf der zwischen Desinformation und Meinungsfreiheit abgewogen wurde. Der DSA-Entwurf der EU-Kommission sieht eine solche Erklärung einzelner Löschentscheidungen ebenso wie regelmäßige Transparenzberichte vor.
- Meldungen sollten nicht rein algorithmenbasiert und damit automatisiert bearbeitet werden, sondern Menschen sollten in den Prozess involviert sein. Ein ‚human-in-the-loop‘-Verfahren bestärkt einen institutionalisierten Löschvorgang von Deepfakes und kann die Entscheidungen, wann Deepfakes gelöscht werden, ethisch und sozial wünschenswert ausgestalten. Der DSA-Entwurf sieht eine solche menschliche Involvierung vor.
- Die Moderation und mögliche Löschung von schädigenden Deepfakes sollte ‚outgesourct‘ werden und nicht durch die Plattformen selbst erfolgen. Wichtig ist es hierbei, dass eine Zusammensetzung der Kooperationspartner*innen objektiv und unabhängig erfolgt und ihre Arbeit transparent und für Externe nachvollziehbar ist.

Plattformpolicies (Mediensebstregulierung, Unternehmensethik)

- Spezialisierte Plattformen für pornografische Deepfakes sollten verboten werden, da diese Materialien in der Regel nicht einvernehmlich produziert werden. Deepfakes auf weiteren pornografischen Plattformen sollten verboten und das Schlagwort ‚Deepfake(s)‘ und verwandte Schlagwörter hier deaktiviert werden, um die Verbreitung von nichteinvernehmlichen Deepfake-Videos zu reduzieren.
- Die Selbstregulierung von Onlinediensten (sozialen Medien, Messengern) ist eine wichtige und vielversprechende Art der Governance von Deepfakes. Plattformen können die Effektivität bösartiger Deepfake-Kampagnen entscheidend abmildern und verfügen im öffentlichen Diskurs über große Macht. Ihre Selbstregulierungsbemühungen sollten unterstützt werden. Besonders positiv sind dabei partizipative Ansätze.
- Wenn Onlinedienste als ‚Watchdogs‘ über bereits publizierte Inhalte fungieren, kann sichergestellt werden, dass die Beurteilung von Inhalten nicht im Vorfeld („ex ante“) und damit nicht in Form eines Generalverdachts geschehen. Dies erschwert es Onlinediensten, auf undemokratische Art und Weise die freie Rede einzuschränken. Durch Beobachtung und Beschwerden von Nutzer*innen kann ein Prozess der Selbstregulierung in Gang gesetzt werden.
- Jedoch sollten Plattformen ihre Detektion und ihren Umgang mit Deepfakes öffentlicher als bisher kommunizieren, um ihre Transparenz und Rechenschaftspflicht zu erhöhen. Onlinedienste sollten Nutzer*innen informieren, warum ihre Inhalte gelöscht oder stummgeschaltet werden, damit sie die Entscheidung anfechten können.
- In einer demokratischen Gesellschaft ist es notwendig, die Selbstregulierung von Onlinediensten durch ethisch wünschenswerte und konkrete Governancemechanismen (insbesondere rechtliche Regulierungen) zu ergänzen, die Onlinediensten einen Rahmen für ihre Selbstregulierung bieten und den schädlichen Nutzungen von Deepfakes entgegenwirken.

Technische Normen und Zertifizierung

- Industriell geprägte technische Normen (und deren Zertifizierung durch technische Prüforganisationen) könnten die Dynamik der Deepfake-(Detektions-)Entwicklung ausbremsen.
- Dennoch lassen sich aus Debatten zur Normierung von KI wichtige Erkenntnisse für Deepfakes ableiten. So sollte auch ohne technische Normung auf die Qualität der Trainingsdaten geachtet werden. Zudem sollten Detektionsentscheidungen für den

Menschen zu einem gewissen Grad nachvollziehbar sein, etwa durch Transparenz in Bezug auf die Trainingsdaten und Fehlerwahrscheinlichkeiten.

KI-Ethikleitlinien und Zertifizierung

- Eine Zertifizierung ist ein stark regulatorischer Eingriff, der Innovationen und Dynamik im Bereich Detektion hemmen könnte. Die Dynamik ist angesichts des ‚Katz-und-Maus‘-Spiels mit der Deepfake-Entwicklung jedoch entscheidend.
- Dennoch lassen sich aus Debatten zu KI-Ethikleitlinien und Zertifizierung wichtige Erkenntnisse für Deepfakes ableiten. Aufgrund der potenziell hohen Kritikalität von Deepfake-Detektionsalgorithmen sollten sie ein gewisses Maß an Transparenz, Nachvollziehbarkeit und Erklärbarkeit für den Menschen aufweisen, einschließlich Informationen über Trainingsdaten, Kriterien und ‚Fehlerwahrscheinlichkeiten‘ der getroffenen Aussagen. Zudem müssen die menschliche Autonomie und Entscheidungshoheit bedacht werden.

Deepfake Detektion

- Detektionssysteme können Anhaltspunkte für die Echtheit von Medien liefern, Schäden für Rezipient*innen verringern und das Vertrauen in die Beweiskraft von Inhalten erhöhen.
- Ethische Grundsätze wie Transparenz, Erklärbarkeit, menschliche Autonomie und Gerechtigkeit sollten bei dem Einsatz von Detektionen beachtet werden. Angesichts der möglichen Verbesserung von Deepfakes aufgrund von Detektionsfortschritten, ist Transparenz etwa gegenüber einer unabhängigen Behörde oder Prüforganisation, die Geheimhaltungspflichten unterliegt, geboten. Zudem muss auf die Qualität der Trainingsdaten und möglichen Bias geachtet werden.
- Detektionssysteme werden nie perfekt sein und können Deepfakes und deren Auswirkungen nicht gänzlich eliminieren. Deshalb ist es nötig, die Ergebnisse von Detektionssystemen in menschliche Entscheidungen einzubinden. Darüber hinaus muss die Deepfake-Detektion von weiteren Governancemechanismen flankiert werden.

Forschungsförderung

- Eine stärkere Förderung der Detektionsforschung in Deutschland insbesondere von staatlicher Seite kann ein Gegengewicht zur privaten und insbesondere vom Militär geförderten US-amerikanischen Technologieentwicklung bilden.

- Bei der Entscheidung über Förderung sollte kritisch reflektiert werden, welche Auswirkungen diese und bestehende Programme auf die politische Nachrichtenübermittlung und Meinungsbildung haben, inwieweit Machtkonstellationen dadurch verändert werden und welche Auswirkungen dies auf die Gesellschaft hat.
- Die Förderbekanntmachung des BMBF ‚Erkennen und Bekämpfung von digitalen Desinformationskampagnen‘ vom Juli 2020 sowie Förderprogramme der EU sind Schritte in die richtige Richtung. Interdisziplinäre Verbände, die im Rahmen gefördert werden, sollten geistes- und sozialwissenschaftliche Forschung mit einbinden. Die Förderprogramme sollten zudem durch weitere gezielte Fördermaßnahmen zur Prävention und Entkräftung von Desinformation und für einen ethisch und sozial wünschenswerten Umgang mit Deepfakes flankiert werden.

Authentifizierung

- Eine freiwillige Nutzung von Authentifizierungssystemen könnte das Vertrauen in bestimmte Medien erhöhen. Gleichzeitig ist es zwingend nötig, dass eine Signierung von unterschiedlichen und unabhängigen Prüfinstanzen kritisch untersucht wird.
- Authentifizierungssystemen können nicht-intendierte Folgen haben, etwa indem sie die Vertrauenswürdigkeit weniger gefälschter Dateien erhöhen. Die Sicherheit von Journalist*innen und Oppositionellen kann durch Authentifizierungssysteme gefährdet werden, zum Beispiel, wenn Journalist*innen und Oppositionelle verdeckte Informationen oder Menschenrechtsverletzungen enthüllen. Eine Authentifizierung gefährdet ihre Anonymität und kann zu Verfolgung führen. Daher ist der verstärkte und insbesondere verpflichtende Einsatz von Authentifizierung ethisch höchst ambivalent. Institutionen wie Gerichte und Versicherungen, aber auch soziale Medien sollten eine Authentifizierung von Bild- und Audiomaterial daher nicht zur generellen Anforderung machen, wenn Medien als ‚Beweise‘ vorgelegt oder etwa auf Plattformen gepostet werden.

Förderung professioneller Faktenprüfung

- Eine Pluralisierung der Akteur*innen im Bereich der Faktenprüfung sollte gefördert werden, um einer parteinahen Affiliation entgegenzuwirken und eine Monopolisierung und interessensgeleiteten Beeinflussung der Faktenprüfung entgegenzuwirken.

Selbst- und Opferschutz

- Ein umfassender Selbstschutz würde große Datensparsamkeit im virtuellen Raum erfordern, um die Menge an potenziellen Daten, die von Dritten abgegriffen und missbraucht werden können, zu minimieren. Dies ist aber kaum umsetzbar. Aus ethischer Perspektive birgt die Forderung nach ‚Selbstschutz‘ zudem die Gefahr einer unzulässigen Verantwortungsabgabe an (häufig besonders vulnerable) Individuen.
- Daher müssen ergänzend zum Selbstschutz von staatlicher Seite präventive Maßnahmen gefördert werden. Dazu gehört die Implementierung einer integrativen Medienkompetenz in das Bildungssystem, um ein gesamtgesellschaftliches Bewusstsein für das Schädigungspotential durch zirkulierende persönliche Daten zu schaffen.
- Es muss zudem ein institutioneller Opferschutz implementiert werden, der durch niederschwellige Beratungsangebote eine rechtliche, soziale und psychologische Unterstützung für Opfer von schädigenden Deepfakes bietet.

Berufsethik

- Ethische Folgenabschätzung sollte fester Bestandteil von Studiengängen wie Informationstechnik, Künstliche Intelligenz oder Ingenieurwissenschaften sein, um die Kompetenzen einer ethischen Reflexion des eigenen Forschens im Feld von Technikinnovation schon frühzeitig zu erwerben.
- In Firmen sollten organisatorische Strukturen (einschließlich des Schutzes von Whistleblowern) implementiert werden, die Raum für ethische Fragestellungen institutionalisieren.
- Private Entwickler*innen sollten stärker in den Diskurs um das Schädigungspotential durch Deepfakes eingebunden werden. Dies könnte über Foren wie GitHub und Reddit erzielt werden. Aufklärungskampagnen und kostenlose Schulungen zum Missbrauchspotential von Deepfakes könnten hier gestreut werden. Die Durchführung von ‚Hackatons‘ zu ethisch wünschenswerten Entwicklungen kann ebenfalls dazu beitragen. Auch zivilgesellschaftliche Angebote für Entwickler*innen sollten gestärkt werden.

Bildung und Medienkompetenz

- Besonders wichtig ist bei Deepfakes Bildung und Aufklärungsarbeit für Medienkompetenz der breiten Öffentlichkeit. Nicht alles kann durch Gesetze und technische Detektion gelöst werden; das kritische Hinterfragen von Nachrichten und Quellen ist zentral, ebenso wie der verantwortungsvolle Umgang mit sozialen Medien. Eine systematische Medienbildung sollte daher stark institutionell gefördert werden.
- Medienkompetenz sollte ein integraler Bestandteil der Ausbildung von Lehrkräften sein, damit sie dieses Wissen in der Schulbildung integrieren können.
- Bildungsträger und zivilgesellschaftliche Organisationen sollten institutionelle und projektbezogene Förderung erhalten, um Informationskampagnen und Aufklärungsmaterial zielgruppengerecht zu konzipieren und durchzuführen.
- Bildungsmaterialien für Medienkompetenz sollten verbreitet werden, die sowohl im formellen wie informellen Bildungssektor und auch in der Erwachsenenbildung einsatzfähig sind.

Bildquellen:

Memed_Nurrohmadi auf Pixabay (S. 1, 9)
Ridwan Hardjowibowo auf Pixabay (S. 23)
Teguhjati Pras auf Pixabay (S. 31)

EBERHARD KARLS
UNIVERSITÄT
TÜBINGEN



**Internationales Zentrum
für Ethik in den
Wissenschaften (IZEW)**

© 2021

Internationales Zentrum für Ethik in den Wissenschaften (IZEW)
Eberhard Karls Universität Tübingen
Wilhelmstr. 19, 72074 Tübingen

Telefon: +49 / 7071 / 29 77981
Telefax: +49 / 7071 / 29 5255

izew@uni-tuebingen.de
www.izew.uni-tuebingen.de

ISBN: 978-3-935933-19-3