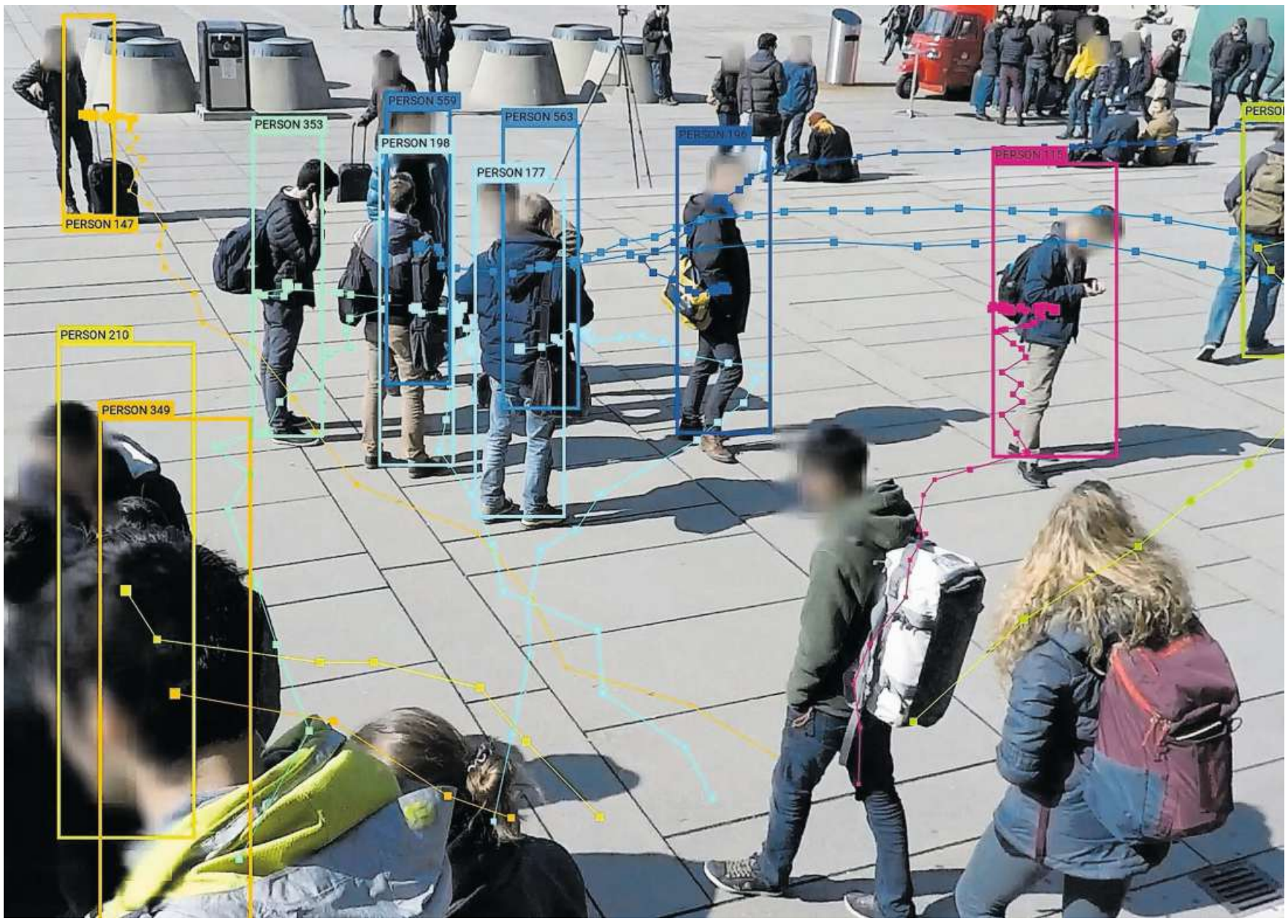



*Vor der ETH im Jahr 2016 gemachte Videoaufnahmen sind frei abrufbar.* VIDEOSTILL WILDTRACK MULTI-CAMERA PERSON DATASET (CHAVDAROVA ET AL.) IDIAP / EPFL / ETHZ; BEARBEITUNG ADAM HARVEY

«Wahrscheinlich wären manche Leute sogar gerne Teil einer Datenbank. Aber sie sollten diese Wahl bewusst treffen können.»

**Adam Harvey**
Wissenschafter und IT-Künstler

# Unsere Bilder tragen zur Überwachung bei

*Online gepostete Fotos und Videos von Passanten landen in Datenbanken. Sie dienen der Optimierung von Gesichtserkennung und anderer künstlicher Intelligenz.*

JUDITH KORMANN

—

Marcus Pfister bezeichnet sich selbst als Dinosaurier, jedenfalls dann, wenn es um neue Technologien geht. Entsprechend spärlich ist der Online-Auftritt des Berner Kinderbuchautors: eine simple, etwas aus der Zeit gefallene Website, eine Facebook-Site, deren letzter Eintrag von 2011 stammt, und keine zehn Bilder auf Instagram. Umso erstaunter war Pfister, als er erfuhr, dass sich Fotos von ihm, die im Netz stehen, in einer Datenbank befinden, mit der Software zur Gesichtserkennung getestet und trainiert wurde. «Ich finde das seltsam», sagt er am Telefon, «es würde mich sehr wundern, wenn man mit den wenigen Bildern etwas anfangen kann.»

Doch Pfisters Gesicht ist eines von rund 100 000, die sich in der Bilddatenbank mit dem Namen MS Celeb befinden. Erstellt hat die Sammlung der Softwarekonzern Microsoft. Sie enthält mehr als acht Millionen Fotos. Die Bilder wurden aus dem Internet abgegriffen, um eine Technologie zu verbessern, mit der Smartphone-Besitzer ihr Handy entsperren, die Sicherheitsbehörden zur Fahndung von Kriminellen einsetzen und autoritäre Regime zur Überwachung der Bevölkerung.

## Datensammeln wie Clearview

Gesichtserkennung und andere Anwendungen der künstlichen Intelligenz (KI) sind auf dem Vormarsch und werfen viele Fragen auf. Sie können das Leben erleichtern, bieten aber auch Potenzial für Missbrauch. Bei der Debatte, wie und ob die Technologien eingesetzt werden sollen, gerät eine Frage häufig in den Hintergrund: diejenige, was nötig ist, um KI voranzutreiben – unsere Gesichter, Stimmen und Körper.

Dass Datenbanken, deren Bilder aus dem Internet stammen, dazu dienen, Gesichtserkennung zu trainieren, hat zuletzt der Skandal um Clearview gezeigt. Die kleine Firma aus den USA hatte laut der «New York Times» mehr als drei Milliarden Fotos aus Plattformen wie Facebook oder Youtube abgesaugt, um ihre Software zu optimieren. Diese soll im Stande sein, Millionen von Menschen innert weniger Sekunden zu identifizieren.

Die Art, wie Clearview seine Software trainiert hat, ist keine Ausnahme. Damit Algorithmen besser darin werden, Gesichter zu erkennen und sie Personen zuzuordnen, füttern Forscher sie mit Hunderttausenden von Fotos. Je unterschiedlicher die Personen darauf, je natürlicher und ungestellter die Aufnahmen; je näher an der Situation, in der die Technik zum Einsatz kommt, desto besser. Öffentlich zugängliche Fotosammlungen wie MS Celeb finden ihren Weg in Forschungsinstitute von den USA über Europa bis nach China. In vielen Fällen stammen die Bilder darin aus dem Internet. Die betroffenen Personen haben davon meist keine Ahnung.

Marcus Pfister ist einer von rund 200 Schweizern, die sich in der Datenbank MS Celeb befinden. Unter ihnen sind Musiker, Sportler, Politiker, Architekten oder Schriftsteller. Von Martin Suter bis zu Simonetta Sommaruga und Doris Leuthard. Offiziell umfasst MS Celeb nur «Celebrities», deren Bilder weniger geschützt sind. Laut Adam Harvey, dem Wissenschafter und IT-Künstler, der die Existenz der Bilddatenbank publik gemacht hat, wurde der Begriff aber weit gedehnt.

In der Sammlung tauchten auch Aktivisten, Journalisten oder Blogger auf, also schlicht Personen, die einen Online-Auftritt haben. Von jenen Schweizern, die wir neben Pfister kontaktiert haben, wusste niemand, dass sie Teil einer Datenbank sind, mit der Überwachungstechnik trainiert wurde. Sie wurden nicht um ihr Einverständnis gefragt. «Wenn ich sagen könnte, ‹meine Bilder lieber nicht›, dann würde ich das tun», sagt der Kinderbuchautor, «aber wie soll man so etwas kontrollieren?»

Microsoft hatte MS Celeb 2016 ins Internet und damit Wissenschaftern weltweit zur Verfügung gestellt. Wie Harveys Recherchen und Research-Papers zeigen, wurde die Sammlung genutzt von Tech-Riesen wie IBM und Hitachi. Auch die chinesischen Firmen Megvii und SenseTime nutzten sie. Deren Software verwenden Chinas Behörden laut Recherchen der «New York Times», um die Minderheit der Uiguren zu überwachen. Vergangenen Sommer, nachdem die «Financial Times» über MS Celeb berichtet hatte, nahm Microsoft die Datenbank vom Netz. Verschwunden ist die Sammlung damit aber nicht. Wer sie heruntergeladen hat, kann sie weiter nutzen. Und über akademische Filesharing-Programme lässt sie sich noch immer finden.

## Jagd nach «wilden Aufnahmen»

«Das ist das Problem mit diesen Datenbanken. Wenn du einmal in einer auftauchst, bekommst du deine Bilder da nie wieder raus», sagt Adam Harvey. Er hat ein Dutzend solcher Sammlungen analysiert. Harvey ist Amerikaner. Heute lebt der 38-Jährige in Berlin. Seit 2010 beschäftigt er sich mit Bilddatenbanken, die ohne explizite Zustimmung der Personen darin erstellt werden. Mit seinem Projekt Megapixels untersucht er, woher die Bilder kommen und wofür sie verwendet werden. «Leute sollen verstehen, was mit ihren Fotos im Netz passieren kann», sagt er in einem Telefongespräch über die verschlüsselte App Signal. Präzise Zählungen dazu, wie viele solche Bilddatenbanken heute im Umlauf sind, gibt es nicht. Harvey geht von 300 bis 400 aus. Bei einem Teil davon handle es sich um Kopien bereits existierender Sammlungen.

Bevor er begann, das Netz nach Foto- und Videosammlungen zu durchforsten, hatte Harvey Ingenieurwissenschaften und Fotografie studiert. Dann entschied er sich, einen anderen Weg einzuschlagen. «Mir wurde klar, dass Fotografie nicht nur Kunst ist, sondern auch ein Werkzeug, um Informationen über eine Person zu sammeln», sagt er, «wenn wir Bilder online stellen, werden diese zu Instrumenten für Überwachung.»

In den 1990er Jahren erstellten Wissenschafter Datenbanken zur Gesichtserkennung noch anhand von Studioaufnahmen. Doch bald wurde ihnen klar, dass sich die Technik damit nur schlecht trainieren lässt. Die Fotos waren zu gut ausgeleuchtet und zu gestellt. Was man brauchte, waren Bilder in natürlichen Situationen. Forscher sprachen von Aufnahmen «in the wild». 2007 veröffentlichte die University of Massachusetts eine der ersten grossen Bilddatenbanken. Sie trägt den Namen «Labeled Faces in the Wild». Die Fotos darin wurden aus dem Internet von Nachrichtensites gezogen und zeigten Personen in verschiedenen Situationen, mit unterschiedlichen Hintergründen und Lichtverhältnissen.

## Für die Forschung «essenziell»

Mittlerweile sei es üblich, Fotosammlungen mit Bildern aus dem Internet zu erstellen, sagt Harvey. Mit einer sogenannten Scraper-Software würden Bilder von Suchmaschinen, Foto- und Videoplattformen abgegriffen. Nicht nur sogenannte «Celebrities» sind betroffen, sondern auch völlig unbekannte Personen. Mindestens eine Sammlung enthält Fotos, die Nutzer auf die Plattform Flickr hochgeladen haben – Kinderfotos inklusive. Die Urheber hatten die Nutzung der Bilder durch eine Creative-Commons-Lizenz bewilligt. «Dabei dachte wohl kaum jemand daran, was das für Folgen haben kann», denkt Harvey. Andere Datensammlungen wurden aus Video-Aufnahmen erstellt, etwa von der Überwachungskamera eines Cafés.

Für Forscher stellen öffentliche Datenbanken eine wichtige Grundlage dar, ob bei Gesichtserkennung oder in anderen Bereichen der KI. «Sie sind essenziell. Die Wissenschaft ist davon abhängig», sagt Luc Van Gool, der das Computer Vision Lab der ETH Zürich

leitet. Denn die Erstellung und Aufbereitung einer eigenen Datenbank ist aufwendig und teuer. Ohne den Open-Data-Ansatz würde man das Feld vollends grossen Konzernen wie Facebook und Google überlassen. «Diese sitzen auf Bergen von Daten und haben einen gewaltigen Vorteil.»

Wo öffentlich zugängliche Datenbanken landen und wie sie genutzt werden, lässt sich aber schwer kontrollieren. Harveys Recherchen zeigen: Auch Forscher von Firmen, die ihre Technik militärischen Einrichtungen zur Verfügung stellen, verwenden sie. Sind sich die betroffenen Personen nicht bewusst, dass ihre Aufnahmen gesammelt und zu solchen Zwecken verwendet werden, wirft das ethische Fragen auf. «Wahrscheinlich wären manche Leute sogar gerne Teil einer Datenbank, denn dadurch werden sie in der Forschung repräsentiert», sagt Harvey. «Aber sie sollten diese Wahl bewusst treffen können und die möglichen Verwendungszwecke kennen.»

### KI-Datenbanken auch in Zürich

Viele KI-Datenbanken kommen aus den USA. Doch sie entstehen auch anderswo, zum Beispiel in der Schweiz. Bei seinen Recherchen ist Harvey auf eine Video-Sammlung gestossen, die in Zürich erstellt wurde, auf der Polyterrasse der ETH. Sie ist nicht für Gesichtserkennung ausgerichtet, sondern für Algorithmen, die erkennen sollen, ob sich Personen im Sichtfeld einer Kamera befinden. Bei dem Projekt handelt es sich um eine Zusammenarbeit der École Polytechnique Fédérale de Lausanne (EPFL), der ETH Zürich und der Forschungseinrichtung Idiap. Es folgt wie viele andere dem Prinzip «in the wild» – also einer so natürlichen Erhebung wie möglich. Und es wirft die Frage auf, ob sich die erfassten Personen darüber im Klaren waren, was mit ihren Aufnahmen geschieht.

2016 positionierten Forscher vor dem Hinterausgang der ETH sieben Kameras und filmten Studierende, Lehrpersonal und Passanten, die durch das Feld liefen. Die Aufnahmen zeigen das Treiben auf der Polyterrasse zeitgleich aus verschiedenen Winkeln: Mehrheitlich junge Leute gehen an den Kameras vorbei, ihre Gesichter sind gut erkennbar. Auch der Ton wird aufgenommen. Viele der Anwesenden scheinen der Aufzeichnung keine grosse Beachtung zu schenken. Manche blicken direkt in eines der Objektive. Ein junger Mann winkt, ein anderer zeigt den Mittelfinger.

Die Videos können auf der Website der EPFL aufgerufen werden. Zwar ist dort festgehalten, dass Dritte sie nur zu Forschungszwecken nutzen dürfen. Wer will, kann sie aber ohne Einschränkung herunterladen. Die Datenbank mit dem Namen «Wildtrack» soll Methoden der Personenerkennung trainieren und testen. Die Anwendungen sind vielfältig. Die Technik kommt laut den Forschern hinter Wildtrack zum Beispiel bei selbstfahrenden Autos zum Einsatz oder, «um Menschen in einer Menge zu zählen».

Die Videos sind aber auch für Technologie nützlich, mit der man eine Person über mehrere Kameras als dieselbe erkennen und so verfolgen kann. Forscher von Microsoft und Wormpex, der Forschungsabteilung einer chinesischen Kette für kassenlose Nachbarschaftsläden, erklärten in einem Paper, sie würden Technologie, die sie mit «Wildtrack» und einer weiteren Datenbank getestet hätten, für diesen Zweck weiterentwickeln. Wissenschafter der chinesischen Nanjing University of Aeronautics and Astronautics und der britischen University of Leicester zitierten die Datensammlung in einem Dokument, bei dem es um die Weiterentwicklung von Drohnen geht, die Menschen aus einer gewissen Entfernung erkennen sollen.

Wenn die Datensammlung in diesem Zusammenhang eingesetzt worden sei, würde ihn das erstaunen, erklärt François Fleuret, Wildtrack-Projekt-Verantwortlicher am Institut Idiap. Da die Aufnahmewinkel nicht dieselben seien, sei die Datenbank für diesen Zweck schlecht geeignet. Sie sei definitiv nicht für Überwachungstechnologien erstellt worden.

> Dritte dürfen Videoaufnahmen von der Polyterrasse der ETH in Zürich für Forschungszwecke nutzen. Wo sie landen, ist nicht steuerbar.

Studierende und Dozenten der ETH wurden laut der EPFL nicht im Vorfeld über die Aufzeichnung informiert. Man habe aber die nötigen Vorkehrungen getroffen, erklärt die Pressestelle in Lausanne: Die Bewilligungsstelle der ETH Zürich habe eine Bewilligung erteilt (diese stellt den Datenschutz in die Verantwortung des Antragstellers, in diesem Fall laut ETH ein Angestellter von Idiap). Die Kameras seien gut sichtbar aufgestellt gewesen. Wer nicht durch das Aufnahmefeld gehen wollte, hätte leicht ausweichen können. Schilder an den Kameras hätten angegeben, dass ein Dreh stattfand – zu Forschungszwecken. Unter einer E-Mail-Adresse hätten gefilmte Personen auch die Löschung ihrer Sequenz fordern können. Das habe niemand getan.

Wofür genau ihre Aufnahmen verwendet werden, wurde den gefilmten Personen, ausser auf Nachfrage, aber offensichtlich nicht mitgeteilt. Noch konnten sie damit rechnen, dass die Videos anschliessend im Internet stehen würden, wo sie – trotz dem Hinweis – nicht nur Forschungsinstitute, sondern auch andere Nutzer herunterladen können.

### «Das ist für mich nicht okay»

Roland Mathys, Rechtsanwalt in Zürich mit Schwerpunkt Datenschutz, der die Videos einsehen konnte, sieht in dem Vorgehen «zumindest Verbesserungspotenzial». Man hätte Studierende und Lehrpersonal detaillierter und bereits im Vorfeld informieren können. Die Formulierung, dass «zu Forschungszwecken» gefilmt würde, hält er für zu allgemein. Grundsätzlich müsse es Personen möglich sein, zu erkennen, dass gefilmt werde, bevor sie den Aufnahmebereich beträten. Dass das in diesem Fall so gewesen sei, bezweifle er. Als «problematisch» stuft er ein, dass die Videos, auf denen die Personen gut erkennbar sind, offen ins Netz gestellt wurden.

«Hätten wir die Gesichter unkenntlich gemacht, hätte das die Übungsdaten verfälscht», erklärt François Fleuret. Luc Van Gool von der ETH, der Teil des Wildtrack-Forscherteams war, sieht das gleich. Er räumt aber ein, dass man Datenbanken hinter eine Schranke stellen könnte, so dass Forscher, die diese nutzen wollten, darum anfragen müssten.

Studierende der ETH waren von den Aufnahmen überrascht. Sehr grosses Kopfzerbrechen scheinen sie ihnen allerdings nicht zu bereiten. «Ich kann mich dunkel an die Kameras erinnern», sagt Tierry Hörmann. Der 24-jährige Informatikstudent ist Vorsitzender des Studierendenverbandes der ETH. Zu welchem Zweck damals gefilmt wurde, wusste er nicht. «Im Sinne der Forschung finde ich persönlich das aber in Ordnung», sagt er. Sein Kollege Florian Moser ist kritischer. «Ich würde mich so einer Aufzeichnung gerne entziehen.» Er wolle nicht zu einer Technik beitragen, bei der man nicht ausschliessen könne, dass sie zur Überwachung eingesetzt werde. «Auch dass die Videos im Internet frei abrufbar sind, ist für mich nicht okay.» Aber, fügt er hinzu, auch er gewichte das Interesse der Forschung höher. Und andere Unternehmen und Institute gingen beim Datensammeln im Internet um einiges schlimmer vor. Schliesslich würden unsere Informationen heute ohnehin überall abgegriffen.

Mitarbeit: Stefan Betschon, Parijat Ghoshal, Barnaby Skinner

# Der Ceneritunnel lässt Subventionen fliessen

*Schweiz sucht Neat-Anschluss via Frankreich*

HELMUT STALDER

Wenn es um die Verlagerung des Güterverkehrs von der Strasse auf die Schiene geht, sitzt dem Parlament das Geld locker in der Tasche. Die bisherigen Massnahmen, um die Lastwagen auf die Bahn zu bringen, wirkten, aber sie reichten nicht aus, befand der Nationalrat bei der Beratung des Verlagerungsberichts. 2018 fuhren zwar 13 000 Lastwagen weniger durch die Alpen, nämlich 941 000. Ohne den Bau der Neat mit den Basistunneln am Lötschberg und am Gotthard, ohne die LSVA und ohne Verbilligungen wären es heute schätzungsweise 1,75 Millionen Camions. Aber trotz dem Rückgang liegt das Verlagerungsziel von 650 000 Lastwagen in weiter Ferne.

Mit den bisherigen Massnahmen allein könne das Ziel nicht erreicht werden, stellte der Bundesrat fest. Sogar die Eröffnung des Ceneri-Basistunnels im Herbst und die durchgängige Öffnung des Vier-Meter-Korridors auf der Gotthardachse würden nicht reichen. Die Gründe hierfür ortet der Bundesrat in vier Bereichen: Erstens ist Deutschland mit dem Ausbau der Zulaufstrecken am Rhein massiv in Verzug. Zweitens fehlt die Interoperabilität im ganzen Korridor. Drittens ist es unmöglich, durchgehend 740 Meter lange Züge zu führen. Viertens sind die Transporte oft unzuverlässig und unpünktlich.

### 385 Millionen zusätzlich

Der Bundesrat beantragte deshalb, den unbegleiteten kombinierten Verkehr stärker zu fördern. Er will den Zahlungsrahmen um 90 Millionen auf 1,765 Milliarden Franken aufstocken und um drei Jahre bis 2026 ausdehnen. Damit stünden zusätzlich 30 Millionen Franken pro Jahr zur Verbilligung des Schienengütertransits zur Verfügung. Doch der Nationalrat zeigte sich spendabler: Fast einmütig erweiterte er den Zahlungsrahmen um 385 Millionen auf 2,06 Milliarden Franken. Zugleich dehnte er die Laufzeit bis 2030 aus, so dass für sieben weitere Jahre je 55 Millionen Franken für «Betriebsabgeltungen» bereitstehen.

Solange die europäischen Zufahrten nicht den Korridor-Standard der Neat erfüllen, ist es absehbar, «dass eine Rückverlagerung auf die Strasse die Folge sein könnte», begründete Kommissionspräsident Kurt Fluri (fdp., Solothurn) die Aufstockung. Sprecher von SP, CVP, Grünen und Grünliberalen betonten, damit werde die Verlagerung gestärkt. Christian Wasserfallen (fdp., Bern) wies auf massive Probleme auf den Zufahrten hin, die zu grossen Defiziten führten. So seien nur 55 Prozent aller Güterzüge weniger als eine Stunde verspätet.

Verkehrsministerin Simonetta Sommaruga plädierte vergeblich für den niedrigeren Betrag und warnte vor Fehlanreizen und Mitnahmeeffekten. «Die ungenügende Qualität wegen fehlender Pünktlichkeit kann man nicht nur mit höheren Betriebsbeiträgen überwinden», sagte sie. Weil die Probleme zum Teil hausgemacht seien, könne man von den Akteuren ebenfalls Anstrengungen verlangen. Schliesslich sei es das Ziel der Subventionen, die Zahl der alpenquerenden Fahrten auf der Strasse zu reduzieren und nicht die Menge im alpenquerenden kombinierten Schienengüterverkehr zu maximieren.

### In Deutschland blockiert

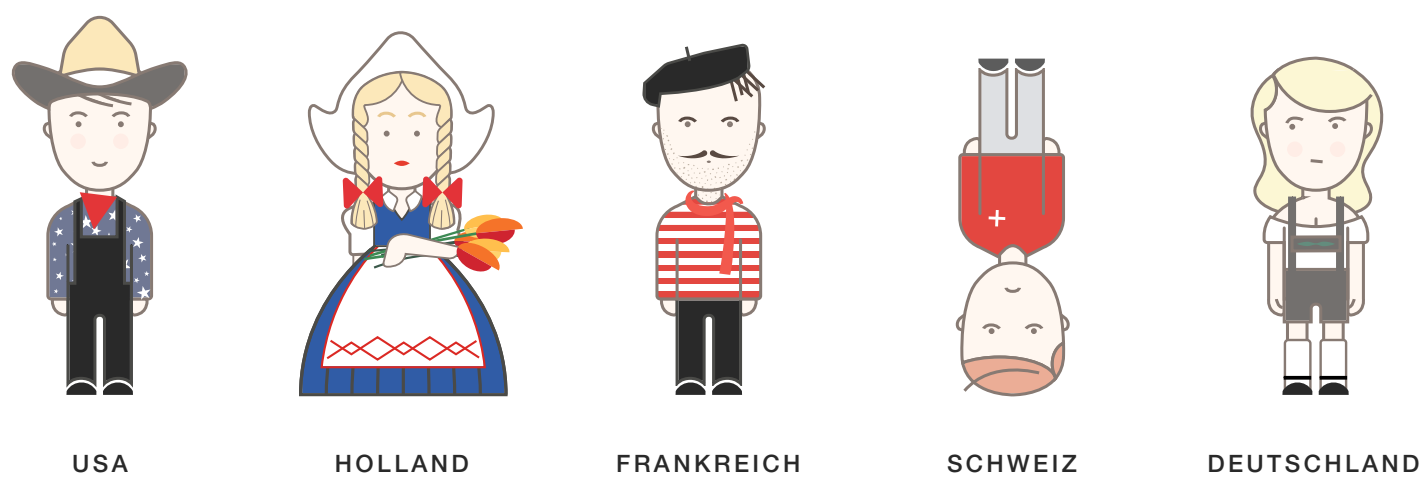
Das grösste Probleme orteten Bundesrat und Nationalrat jedoch bei der Zufahrtsstrecke zur Neat auf der deutschen Rhein-Seite. Der durchgängige viergleisig Ausbau der Strecke zwischen Karlsruhe und Basel, wie ihn die Schweiz und Deutschland 1996 im Staatsvertrag von Lugano vereinbart hatten, ist in Deutschland durch Einsprachen blockiert und soll erst 2040/41 vollendet sein. Deshalb überwies der Nationalrat oppositionslos eine Motion, die den Ausbau der linksrheinischen Strecke in Frankreich verlangt. Der Bundesrat soll mit Frankreich und Belgien einen Staatsvertrag aushandeln.

Sommaruga versprach, alles daranzusetzen, Verbesserungen auf dieser Route zu erreichen. Falsche Hoffnungen dürfe man sich aber nicht machen, denn die linksrheinische Strecke habe für Frankreich keine strategisch vorrangige Bedeutung. «Wir werden zeigen müssen, dass wir allenfalls auch bereit sind, uns finanziell zu engagieren», sagte Sommaruga. Dass die Schweiz Bahnstrecken im Ausland mitfinanziert, ist durchaus üblich, wenn sie in ihrem Interesse liegen. Beiträge leistete die Schweiz zum Beispiel bei den Anschlüssen ans französische Hochgeschwindigkeitsnetz und an die Strecke Zürich–München, aber auch an die Neat-Anschlüsse in Italien via Luino.


**Verkehrte Welt für Schweizer Unternehmen**

Die Unternehmens-Verantwortungs-Initiative und der Gegenvorschlag des Nationalrats wollen einen Alleingang der Schweiz. Sie stellen die rechtliche Welt für Schweizer Unternehmen auf den Kopf.

Es soll neu eine Umkehr der Beweislast gelten. So werden Klagen gegen Schweizer Unternehmen attraktiv. Die Unternehmen müssen vor Gericht beweisen, dass sie alles richtig gemacht haben und werden erpressbar.




**NEIN zur Unternehmens-Verantwortungs-Initiative**

Wirtschaftskomitee «Zusammenarbeit statt Gerichtsprozesse» | www.erpresserische-klagen-nein.ch