

# Traue deinen Augen nicht

Neue Computerprogramme machen die Manipulation von Videos leicht. Das Bild wird Lüge – mit unabsehbaren Folgen **VON ALARD VON KITTLITZ**

**D**enken wir uns zwei Videos: Das erste zeigt Angela Merkel, sie reitet auf einem T. rex und feuert aus einer Bazooka auf Christian Lindner. Im zweiten Video ist ebenfalls Merkel zu sehen, vor einer schlichten grauen Wand, ein Deutschlandfähnchen in der Ecke, die Aufnahme ist wacklig, sie könnte mit dem Handy gefilmt worden sein. Die Kanzlerin sagt in die Kamera: »Wir sind uns doch alle einig, Deutschland muss islamisiert werden.«

einlädt – und voraussichtlich missbraucht werden wird –, dürfte allen Beteiligten längst bewusst sein.

Das alles wird für die Politik einen tiefen Einschnitt bedeuten. Denn jeder Fake dürfte den Wahrheitsbegriff weiter erodieren lassen – und damit jene gemeinsame Basis, auf deren Grundlage politische Entscheidungen getroffen werden. Nehmen wir das Eingangsbeispiel von der Bundeskanzlerin und dem Islam. Angenommen, solch ein Video tauchte in den sozialen Netzwerken auf, gestreut von rechten Gruppen mit der Behauptung, es sei ein »Leak«. Wie viele Likes, Shares und Kommentare würde es generieren? Wie viele Stimmen würde es die CDU kosten? Gewiss würde bald ein Dementi aus dem Kanzleramt auftauchen, das Video würde zur Fälschung erklärt. Aber wie viele der Leute, die ohnehin an den Inhalt des Films glauben wollen, würde das Dementi in ihrer Filterblase erreichen?

Mit jeder Fälschung dürfte sich die Spirale der Verunsicherung ein Stück weiter drehen. Nicht nur Videos verlieren dann an Überzeugungskraft, sondern auch Dementis. Wenn alles Fake sein kann, kommt auch die Behauptung, etwas sei fingiert, immer wohlfeiler daher. Die Folge: Die Wirklichkeit wird noch unübersichtlicher, die Politik opaker.

Schon heute säen Propagandamedien in Form von »alternativen Fakten« systematisch lähmende Zweifel, sei es beim Krieg in der Ukraine oder bei der Behauptung, der Klimawandel sei menschengemacht. Wenn nun auch das Abbild von Politikern zur Marionette von Trollen und Propagandisten verkommt, verschwindet eine zentrale Unterscheidungsmöglichkeit zwischen Information und Desinformation. Das Zeitalter, in dem Wahrheit »festgehalten« werden konnte, in der ein Video Siegel der Authentizität war, wird bald Geschichte sein. Es beginnt eine neue Ära, in der Misstrauen zu jeder Zeit und gegenüber jedem Medium nicht mehr Sache des Paranoikers ist, sondern Teil alltäglicher Medienrezeption.

Dabei stand am Anfang dieser Technologie ein biologischer Zufall: die Mund-Kiefer-Gaumenspalte. 2013 arbeitet der Informatikstudent Justus Thies an der Universität Erlangen an einem medizintechnischen Projekt. Er bastelt anhand von Aufnahmen mit Spezialkameras dreidimensionale, digitale Abbilder von Köpfen solcher Menschen, die wegen einer Mund-Kiefer-Gaumenspalte in Behandlung sind. Sein Ziel ist, anhand von Heilungsverläufen bessere Aussagen über die Chancen anderer Patienten zu machen. Das Thema der dreidimensionalen Konstruktion und Übertragung von Gesichtern lässt Thies nicht mehr los. Nur drei Jahre später, im März 2016, präsentiert er gemeinsam mit Kollegen vom Max-Planck-Institut und der Universität Stanford eine Technologie, die er Face2Face nennt – sie erlaubt es ihm, in Echtzeit die Mimik und Bewegungen eines aufgenommenen Gesichts auf das Gesicht eines Menschen in einer Videoaufnahme zu übertragen.

In einem Erklärvideo, das auch auf YouTube zu finden ist, führt Thies seine Technologie vor. Zu sehen ist Thies, der sich selbst dabei filmt, wie er das Gesicht verzieht – und George W. Bush, auf einem Fernseher neben ihm, der Thies' Bewegungen wiederholt. Das Video von Bush stammt ursprünglich von CNN, seine Mimik wird aber von Thies kontrolliert. Bush guckt nicht mehr, wie er im Original-Interview guckte, sondern so, wie der Informatiker es will. Thies lächelt – Bush lächelt. Thies zieht die Brauen zusammen – Bush zieht die Brauen zusammen. Der Programmierer steuert das Gesicht des Politikers fern. Und der Fake ist nicht erkennbar.

In einer E-Mail an die ZEIT erwähnt Thies die vielen potenziellen Anwendungsfelder seiner Technologie. In Autos könnten Warnsysteme damit das Gesicht und den Wachheitszustand eines Fahrers auch bei schwierigen Lichtverhältnissen kontrollieren, die Entertainment-Branche könnte noch wirklichkeitstreuere Virtual-Reality-Szenarien entwickeln. Doch Thies dürfte wohl auch das Gefahrenpotenzial seiner Technologie bewusst sein, immerhin demonstriert er sie an Politikern. Für die neueste Version von Face2Face, schreibt Thies, bräuchte es schon keine Spezialkameras mehr.

Nun ist die Technik, die hinter Face2Face steckt, kompliziert und nicht gänzlich öffentlich. Die wichtigsten Codes hat Thies für sich behalten. Aber solche Arbeit, so sagt er selbst, ist reproduzierbar. Er verweist auf das jüngst von der Universität Washington veröffentlichte Projekt Synthesizing Obama, für das Wissenschaftler das Gesicht des ehemaligen Präsidenten digitalisierten und es nun reden lassen können. Sie füttern ihr Obama-Programm mit einer Audiodatei, einer Rede zum Beispiel, und das synthetische Gesicht Obamas spricht diese Rede nach. Dass es sich um eine Simulation handelt, ist kaum erkennbar. In einfacherer Form sind ähnliche Mechanismen schon im Umlauf: Russische Programmierer haben kürzlich eine Handy-App präsentiert, mit der die User Bewegungen des eigenen Gesichtes filmen und auf Gesichter in Fotos übertragen können – die Fotos werden zum Leben erweckt. Noch lässt sich damit vor allem Quatsch machen, ein Covergirl Grimassen schneiden lassen zum Beispiel. Doch je ausgefeilter die Technologie wird, desto ernster wären die möglichen Konsequenzen.

Im Bereich der Stimme ist man schon deutlich weiter. Adobe hat mit VoCo vergangenes Jahr den Prototypen eines Programms vorgestellt, das lediglich mit genug Audiomaterial gefüttert werden muss – etwa zwanzig Minuten der Stimme eines Menschen reichen –, um danach die Stimme dieses Menschen extrem realistisch simulieren zu können. Der User kann diese Stimme danach einen beliebigen Text sprechen lassen. Wörter können per Copy-and-

Paste aus Sätzen ausgeschnitten oder in sie hineingebaut werden. Füttert man dieses Programm über 20 Minuten mit der Stimme von Alexander Gauland, lässt sich danach mit Gaulands Stimme spielen wie auf einem Instrument. Die Maschine imitiert seine Intonation, Stimmlage, Pausen.

Das Potenzial der Bild- und Stimmbearbeitung ist gewaltig – und könnte gegen Regierende ebenso gewendet werden wie gegen Bürger. Wie also ließe es sich einhegen?

Relativ einfach könnte man die Verbreitung von manipuliertem Bild- und Tonmaterial als Verleumdung unter Strafe stellen. Deutlich schwieriger ist es, die Schuldigen auszumachen und ihre Schuld zu beweisen. Längst herrscht in der Welt der Informatik ein verzweifelter Wettlauf zwischen jenen, die Informationen manipulieren, und denen, die die Manipulation zurückverfolgen. Der Kampf zwischen Manipulatoren und Forensikern erinnert an den Wettlauf von Hase und Igel: Die Forensiker entdecken im Quellmaterial einer Datei einen Hinweis auf deren Manipulation und macht diesen publik – und der Manipulator bastelt als Antwort an seiner Technologie, um diesen Hinweis in der nächsten Veröffentlichung verschwinden zu lassen. Die Forensikerin hilft dem Manipulator, Lücken zu finden und zu schließen, und arbeitet ihm damit zu. Die Auswertung von Beweismitteln braucht so immer mehr Zeit. Bis die Forensikerin der Öffentlich-

keit erklären kann, ein Video sei ein Fake, hat es im Zweifel längst seine manipulative Wirkung entfaltet.

Entwickler denken solche Folgen bereits mit. Von Adobe heißt es, Audio-Dateien aus VoCo könnten nach einer Veröffentlichung des Programms mit einer Art digitalem Wasserzeichen versehen werden, das die Herkunft der Tondatei aus dem Computer nachweisen würde. Nur: Wer in manipulativer Absicht handelt und technisch gewieft ist, kann wahrscheinlich auch Wasserzeichen austricksen.

Bleibt die Hoffnung, dass die zunehmende Manipulierbarkeit sämtlicher Daten endgültig das Zeitalter der Kryptotechnik einläutet. Jede Datei, die versendet wird, erhielte von ihrem Ersteller dann eine verschlüsselte Signatur, mit der er für ihre Echtheit bürgt. Der Rezipient könnte entscheiden, ob er den Ersteller für glaubwürdig hält oder nicht. Erhielte die ZEIT von ihrer Korrespondentin in Syrien ein Video mit einsehbarer digitaler Signatur, wüsste die Redakteurin im Haus um die Glaubwürdigkeit des Materials, könnte es – wiederum versehen mit der eigenen Signatur der ZEIT – über Kanäle wie Twitter mit jenen teilen, die der ZEIT folgen, und die könnten digital nachvollziehen, dass das Video aus einer geprüften Quelle stammt. Eine Kette des Vertrauens. Der Rest, alles, was nicht aus solchen Ketten stammt, wäre – manipulativer Müll. Spam.

[www.zeit.de/audio](http://www.zeit.de/audio)



So könnte es aussehen, wenn in Zukunft Videos von Politiker-Reden beliebig manipuliert werden können

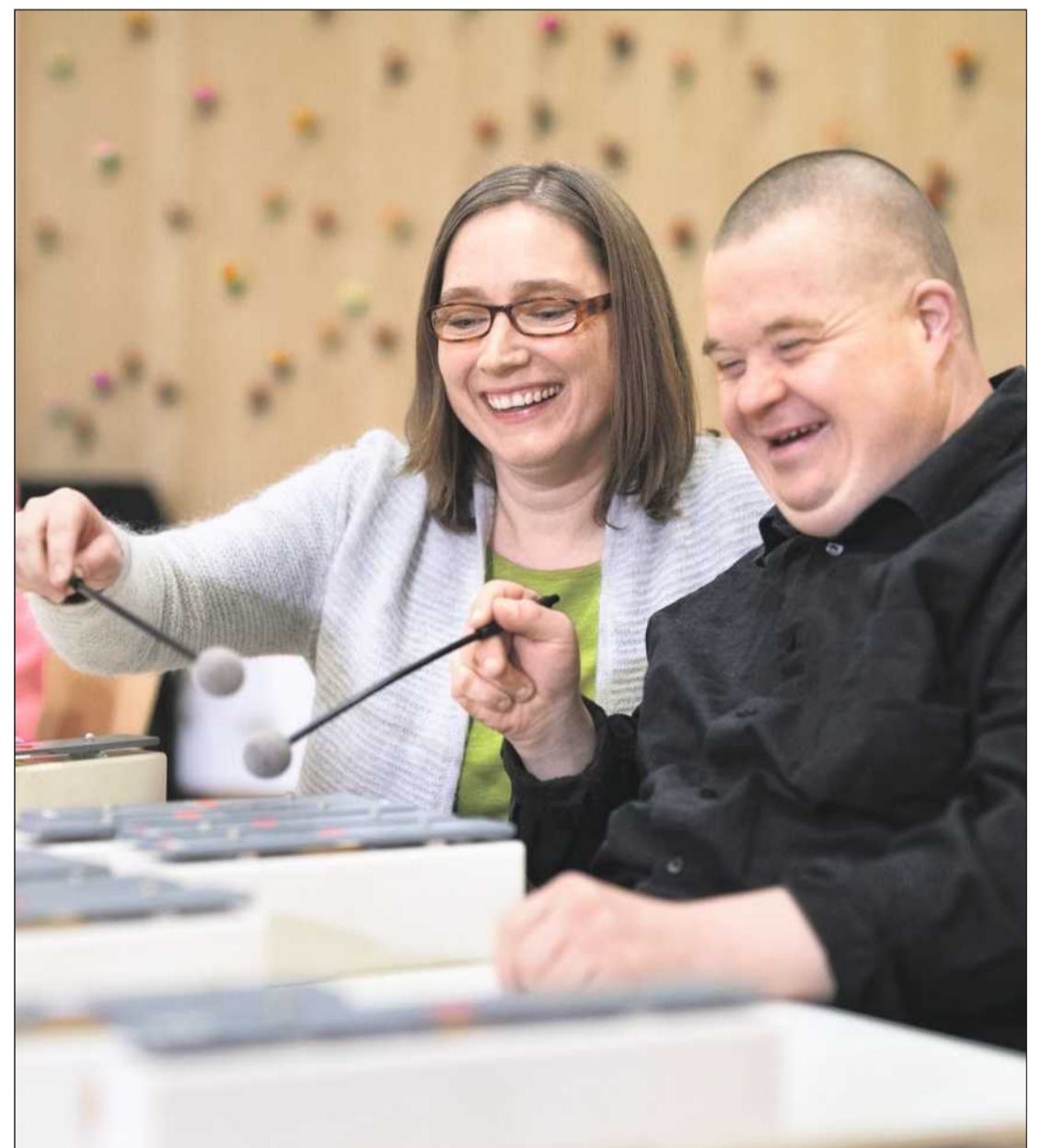
Beide Videos gibt es nicht, sie sind erfunden. Sie sollen hier bloß illustrieren, wie unterschiedlich wir bewegte Bilder rezipieren. Merkel auf dem T. rex – wir wüssten, das kommt aus der Trickfilmkiste. Aber eine Merkel, die ausschaut wie immer, die selbst spricht, vor offizieller Kulisse? Was sie sagt, mag uns stutzig machen. Aber was, wenn an den Bildern keine Manipulation zu erkennen ist?

Fotos betrachten wir spätestens seit der Erfindung von Photoshop mit einiger Vorsicht. Filme dagegen galten bislang als überführend. Wenn ein Video Donald Trump dabei zeigt, wie er einen Behinderten nachhäft, oder Jogi Löw beim Nasepopeln, dann ist für den Zuschauer klar: Das ist so geschehen. In der massenmedialen Demokratie gelten Filmaufnahmen als letzte Bastion gegen den Fake. Diese Bastion aber fällt.

Denn die Technologie, mit der sich Bewegtbild und Ton manipulieren lassen, hat allein im vergangenen Jahr immense Sprünge gemacht. Es ist nun möglich, Köpfe virtuell sprechen zu lassen – ohne dass der Fake noch zu erkennen wäre. Und dank neuer Software braucht es dazu nicht einmal mehr einen Superrechner, ein einfacher Laptop genügt. Die Mimik, die Mundbewegungen von Politikern in Videos lassen sich täuschend echt, bei Bedarf auch live, verändern und steuern. Die Stimme eines Menschen kann vom Rechner inzwischen so überzeugend simuliert werden, dass sie ihm nicht länger allein gehört. In den letzten Monaten demonstrierten die Erfinder, Informatiker und Unternehmen, wie potent ihre Software ist. Dass die Technologie zum Missbrauch


Fotos (M): Elmar Klemmer/VEEN SIMON/hip

ANZEIGE



## Du bist unersetzlich.

♥ Danke

 Bundesministerium für Familie, Senioren, Frauen und Jugend

Mehr als 30 Millionen freiwillig Engagierte machen Deutschland lebenswerter.

Jetzt danken unter [du-bist-unersetzlich.de](https://www.du-bist-unersetzlich.de)

  [#unersetzlich](https://www.du-bist-unersetzlich.de)