

Könnt ihr euren Ohren trauen?

Ihr hört meine Stimme, aber bin das auch tatsächlich ich?

Oder spricht hier mit meiner Stimme eine künstliche Intelligenz,  
ein Klon von meiner Stimme.

Was meint ihr?

Auflösung, ich bin zwischendurch.

Ich bin ein Klo.

Auflösung, ich bin's wirklich.

Meine geklohte Stimme kommt später dazu.

Ich spreche heute mit der NDR-Technologie-Journalistin Svea Eckhardt  
über solche Audiofakes und wie sie hergestellt werden.

Es fängt an mit einem Anruf eines vermeintlichen Firmenchefs  
namens Johannes.

Und es geht um eine Welt, in der wir jetzt die Echtheit von Stimmen  
immer wieder anzweifeln werden.

Ihr hört 11 km der Tagesschau-Podcast.

Ein Thema in aller Tiefe.

Mein Name ist Victoria Michalsack.

Heute ist Dienstag, der 27. Juni.

Es ist Freitag-Nachmittag in England

bei einem mittelgroßen Unternehmen, Energieunternehmen.

Freitag-Nachmittag muss man sich mal vorstellen.

Ich bin schon so ein bisschen in Wochenende und chill laune.

Da bekommt der Leiter der dortigen Zentrale  
einen Anruf von seinem deutschen Konzernchef.

Der heißt Johannes.

Die beiden kennen sich, die haben ab und zu mal miteinander telefoniert.

Der große Konzern Oberboss aus Deutschland,

der spricht Englisch mit so einem leicht deutschen Akzent.

Daher kann man ihn auch ganz gut erkennen.

Der ruft an und sagt, es gibt eine ganz wichtige Rechnung.

Die müssen wir noch begleichen, und zwar an Lifferanten.

Das ist ganz dringend vom Wochenende noch raus.

Bei uns ist es schon spät, es ist Freitag,

die Banken haben schon geschlossen,

bei dir ist noch ne Stunde früher, du musst es dringend machen.

Der Chef aus der Konzernzentrale sichert also zu,

ihr bekommt das Geld, was ihr für diesen Lifferanten auslegt.

Das bekommt ihr am Montag wieder zurück.

Da gibt es quasi eine Konzern-Eigene-Rücküberweisung.

Der Vorstand von diesem Tochterunternehmen sagt,

ich kümmerge mich drum, dann gibt es auch noch eine E-Mail.

Da steht dann die E-Bahn und der Verwendungszweck drin.

Das scheint also alles ganz okay zu sein.

Und überweistern 220.000 Euro.

Okay, und was passiert dann?

Ja, es kommt also dem Montag und es kommt keine Rücküberweisung.

Gut, es kommt sogar noch besser.

Denn dieser Johannes, also dieser Firmschef,  
der Konzernchef aus Deutschland ruft wieder an.

Und erzählt dann irgendwas von einer zweiten Zahlung,  
die jetzt auch noch ganz dringend nötig würde.

Und da fängt es dann bei diesem englischen Firmschef an zu dämmern.

Weil es sollte ja eigentlich heute die Rücküberweisung kommen.

Und warum jetzt eine zweite Überweisung?

Dann schaut er noch mal aufs Telefon, da stimmt dann die Vorwahl nicht.

Das ist eine österreichische Vorwahl und keine deutsche.

Also er legt dann auf, er tätigt die Überweisung nicht.

Er ruft den deutschen Konzernchef an.

Und der beteuert, dass er nie angerufen hat.

Und dass er nie eine Zahlung in Auftrag gegeben hat.

Der hat gar nicht mit dem Johannes gesprochen.

Genau, und erzählt er mir diese Geschichte, Marie-Christine Kraag.

Und es ist immer so, dass wir im Schadenfall versuchen,  
nachzuvollziehen, was genau passiert.

Die ist Global Head of Fidelity bei der Allianz Trade.

Die Allianz Trade ist so ein großer deutscher Rückversicherer.

Und Global Head of Fidelity, das ist Leiterin Betrugsabsicherung.

Also macht die weltweit Wirtschaftskriminalität.

Und da landen solche Fälle,

wenn die Versicherungen, die diese Unternehmen wiederum  
gegen solche Schäden versichern.

Über die Jahre hat sich diese Masche extrem weiterentwickelt.

Am Anfang waren es vor, nämlich E-Mails, mal versucht, reingeworfen.

Und die haben dann eben alle befragt,

um sich diese ganzen E-Mails angeschaut und geschaut.

Was steckte denn mutmaßlich hinter dem falschen Johannes?

Die Versicherung ermittelt also und was finden die raus?

Also die Versicherung geht tatsächlich davon aus,  
dass es eine synthetische Stimme war.

Und zwar erzeugt mit künstlicher Intelligenz.

Also kein Stimme-Imitator, der sich mühsam  
irgendwie die Stimme des deutschen Konzernchefs antrainiert hat.

Sondern das Ganze mit KI erzeugt.

Ja, warum?

Weil die davon ausgehen, das ist einer,

der hat eine ganz eigene Stimmelmelodie, die kannten sich,  
die Stimme war wirklich täuschend ähnlich.

Das ist ein deutschen Akzent, der sicherlich nicht ganz so leicht nachzumachen war, hatte diese Sprachpausen.  
Und deswegen kam die Versicherung,  
nachdem sie also dann verschiedene Zeugen befragt hatte,  
kommen die also zu diesem Schluss.  
Das könnte mit KI gewesen sein.  
Und das würde auch passen zu der Masche.  
Die Masche ist bekannt als Fake-President oder CEO-Fraud,  
wo es immer die Gemeinsamkeit gibt,  
dass sich ein Betrüger ausgibt,  
als ein hochrangiger Entscheidungsträger eines Unternehmens.  
Dann tritt er auf einen Mitarbeiter zu  
und konstruiert eine Geschichte, wie zum Beispiel eine dringende Überweisung,  
und bittet unter diesem Vorwand den Mitarbeiter Geld zu überweisen.  
Der Vorteil bei dieser Masche ist eben,  
dass es von diesen Vorstandsvorsitzenden,  
von denen gibt es eben auch reichlich Material,  
weil die müssen ja auf Pressekonferenzen,  
die machen diese Jahresabschluss-Calls,  
sodass man, wenn man eben viel Stimmmaterial von einer Person hat,  
dann kann man eben dem Computer trainieren,  
dass er diese Stimme täuschend echt nachahmen kann.  
Und da sind natürlich alle Menschen besonders gefährdet,  
von denen es viel Stimmmaterial gibt.  
Also CEOs, Vorstände, wichtige Manager,  
aber auch so Leute wie du, liebe Victoria.  
Auch Podcast-Hos, ja, von, ja,  
meine Stimme ist auch da draußen im Internet vertreten, das stimmt.  
Und deswegen machst du uns jetzt vor,  
wie man einen solchen Stimmklon macht, und zwar von mir.  
Genau, das kommt eben jetzt raus.  
Da gibt es auch sehr viel Stimmmaterial.  
Und das ist auch genau das, was das neue Ronalennetzwerk,  
also was der Computer, was die Technik, die dahinter liegt,  
was die braucht.  
Die braucht nämlich echtes Victoria-Stimmmaterial,  
damit sie dich dann täuschend echt nachahmen kann.  
Und das habe ich natürlich auch mitgebracht.  
Es gibt hier so ein kleines Victoria-Sample-Best-Off.  
Aber das ist schon meine erste Frage.  
Wie groß ist dieser Datensatz, den man füttern muss,  
damit das funktioniert?  
Reicht da eine Folge 11 km oder sind es nur ein paar Sätze  
oder sind es so 100 Folgen?

Das kommt wirklich drauf an, wie gut dein Klon werden soll.  
Also du musst dir so vorstellen, je mehr Stoff der Computer von dir hat,  
desto besser, also ich vermenschliche den jetzt mal ganz schön krass,  
aber desto besser kann das neue Ronalennetzwerk deine Stimme lernen.  
Jetzt hat sich aber die Technik schon so sehr verbessert,  
dass man eben gar nicht mehr so viel braucht.  
Denn bei so kommerziellen Anbietern, und einem von dem machen wir das jetzt auch,  
also ich habe jetzt hier nicht irgendwas neu programmiert  
oder erfunden, sondern einfach nur klick-klick.  
Es gibt wirklich unzählig von diesen Anbietern im Internet.  
Das ist jetzt kein spezielles Programm, ist es teuer?  
Nee, also ich habe das jetzt kostet jetzt ein Euro für einen Monat.  
Gut, also das heißt, es ist wirklich für jedermann machbar.  
Genau, es gibt auch kostenlose Programme,  
da habe ich auch ein bisschen rumexperimentiert,  
die finde ich ein bisschen mechanisch.  
Und jetzt für die Folge wollte ich...  
Jetzt die ein-Euro-Variante, ist dann etwas besser, genau.  
Aber man kann tatsächlich sich auch für 20.000 Euro  
oder für mehrere 1.000 Euro so einen richtig geilen Klon machen lassen.  
Okay, das heißt 20.000 Euro, das klingt jetzt nach viel Geld,  
aber ich denke mir, Menschen mit genug, auch krimineller Energie,  
die zum Beispiel mal so 220.000 Euro irgendwo erbeuten,  
die könnte das ein gutes Werkzeug, ein mächtiges sein, oder?  
Auf jeden Fall, also es gibt ein Gerichtsdokument aus den USA,  
wo mittlerer auch dem, also so einem Betrugsfall  
mit so einem Fake-President-Fake-C.O. nachgegangen sind,  
da ging es um 35 Millionen US-Dollar,  
die die Betrüger damit erbeutet haben mit so einer gefälschten Stimme.  
Also von daher, ja, das lohnt sich.  
Für die Folge ist es nicht ganz so viel.  
Wir gucken jetzt diese Variante an.  
Genau, Stimmsample, ich spiel's einmal kurz an.  
Da reichen eigentlich ein paar Minuten.  
Es ist gerade dieses eine Thema, und das ist gefühlt überall.  
Ja, das ist eine FKM-Folge, die wir vor kurzem veröffentlicht haben.  
Fünf Minuten haben wir da jetzt.  
Ah.  
Und ich gehe jetzt also an diese Software  
und dann drücke ich auf Cloned Voice, das ist ein kleines Pluszeichen,  
und das geht dann einfach blim-blum-drag-drop, ja?  
Mache ich hier Instant Voice-Cloning, schreibe ich hin,  
Victoria's Stimme.  
Und jetzt kommt hier der Witz, das müssen wir einfach mal erwähnen.

Da muss man nämlich so ein Hack hinsetzen.

Und da steht dann, ich bestätige, dass ich die Rechte  
und die Erlaubnis habe, dieses Stimmmaterial hochzuladen.

Gebe ich dich hiermit, kannst du anklicken.

Gut, weil da wird's natürlich auch so ein bisschen ...

Klar.

Weil so Stimme ist ja fast wie ein Fingerabdruck  
und da ist ein biometrisches Merkmal, das ist ganz individuell.

Und das ist jetzt hier im Schrank ...

im Schrank dieser Firma.

Gut, und dann laden wir das also hoch,  
und es ist, ehrlich gesagt, schon fertig,  
und dann klick ich auch schon auf Use, also Benutzung.

Ach, jetzt hast du da so ein Textfeld,  
und jetzt kannst du da eingeben, was du willst.

Genau.

Genau, wir machen mal, kannst du mir bitte 220.000 Euro überweisen?

Kannst du mir bitte 220.000 Euro überweisen?

Okay, ich fand, dass die Stimme so überhaupt nicht nach mir klang,  
diesmal.

Ich mach noch mal Generate, das ist nämlich der Witz auch,  
bei diesen neuronalen Netzen.

Also, wir können das jetzt immer wieder generieren und schauen.  
Da wird immer wieder ein Tick verändern.

Ah, okay.

Kannst du mir bitte 220.000 Euro überweisen?

Schon ein bisschen weicher.

Es klingt anders, und das finde ich interessant.

Es ist, finde ich, bei dem Satz jetzt immer noch nicht so perfekt  
an meiner Stimme dran, aber man kann es halt verändern.

Ich glaube, wenn man da ein bisschen dran rumspielt

und ein bisschen an den Reglern dreht,

könnte man bei einer Version landen,

die je nachdem, wer sie hört, vielleicht schon denkt,

dass zumindest dieser eine Satz von mir ist.

Das könnte sein.

Und wenn man es dann vielleicht noch über ein Telefon hört  
oder über eine Sprachnachricht verschickt,

da könnte man schon ein bisschen Schildluder mit treiben auch.

Hey, Yasmin, ich habe blöderweise das Passwort für die Box vergessen.

Ich laufe gerade zum Studio und muss noch schnell was runterladen.

Hi, Katharina, ich habe gerade das Passwort für die Box vergessen.

Kannst du mir das vielleicht noch mal sagen?

Das konnten wir uns natürlich nicht nehmen lassen.

In Vorbereitung für diese Folge haben wir an einige Kolleginnen und Kollegen bei uns aus dem Team so eine gefakete Sprachnachricht von mir geschickt, um mal zu gucken, ob die darauf reinfallen.

Ich muss mal an meinem Handy den Flugmodus rausmachen.

Und dann gucke ich mal.

Es ist angekommen, aber es hat keiner darauf reagiert.

Vielleicht sind die bei uns nicht.

Wir können ja am Ende von der Sendung noch mal nachschauen.

Das Prinzip von falschen Polizisten am Telefon, das kennen wir auch schon aus einer FKM-Folge.

Es gibt ja auch den Enkeltrick, der funktioniert ähnlich.

Und diese kriminellen Banden, die arbeiten ja schon sehr lange, ganz professionell, nutzen die jetzt auch diese künstliche Intelligenz.

Ja, Enkeltrick, ich habe tatsächlich mal bei einigen LKN, also bei einem Landeskriminalämtern mal nachgefragt, weil es mich natürlich interessiert hat.

So gibt es da, inwieweit ist die Polizei da schon auf dem Plan.

Da muss man aber einmal sagen, da sind wirklich alle Antworten, die ich bekommen habe, so ein Fall ist uns nicht bekannt.

Muss man natürlich auch sagen, dass ich mir vorstelle, bei einem Enkeltrick, dass es dann doch relativ aufwendig ist, sage ich mal, sich die Stimme der Tochter oder der Enkelinn zu besorgen, dann die Festnetznummer der Mutter, der Oma rauszufinden, das ist dann doch recht aufwendig.

Also auch, wenn das Klon der Stimme an sich relativ leicht ist, das Besorgen des Materials,

das ist natürlich bei prominenten Personen, wie bei Vorständen, CEOs, Promis, Podcastern, ist das natürlich viel, viel leichter als jetzt bei Privatpersonen.

Es gibt über Berichte aus den USA.

Also ich denke, je mehr wir von uns im Internet preisgeben, je mehr Insta-Videos wir machen, je mehr Stimmmaterial auch von uns zur Verfügung ist im Netz, dass du eher, kann das natürlich auch für solche Deep Wakes, dann eingesetzt werden.

Die Stimmklone ja nicht nur Kriminelle, die tauchen ja inzwischen oft im Netz auf für ganz unterschiedliche Zwecke, ne?

Ja, so ganz bekannt oder viral ging zum Beispiel ein Fake Drake-Song.

Also das war in aller Munde, Millionfach, geklickt und geguckt worden.

Also sag ich mal, in der Musik, in der Pop-Kultur wird damit experimentiert,

dort tauchen diese synthetischen Stimmen auf.

Ich denke, gerade im Bereich Hörbuch oder auch im Bereich Newsletter, wenn ich ein Blogger bin oder ein Podcast, dann ich gebe noch Newsletter raus und kann dann anbieten, hey, du kannst meinen Newsletter auch einfach anhören mit meiner Stimme.

Also es gibt ganz viele Möglichkeiten, für einen guten, spannenden und nice Newscase für diese synthetischen Stimmen.

Aber es gibt natürlich auch diesen Fall zum Beispiel Desinformation.

You know, people might be surprised to hear me say this, but I actually like Ron DeSantis a lot.

Es ist echt gerade ein ganz neues Video, das ist Hillary Clinton.

Also wir haben ja Wahlkampf in USA.

If Ron DeSantis got installed as president, I'd be fine with that.

Beziehungsweise, das ist nicht Hillary Clinton.

Hier wird nämlich behauptet,

Clinton würde den republikanischen Präsidentschaftsbewerber Ron DeSantis mögen oder sogar unterstützen, ist natürlich Quatsch.

Deswegen gucken auch gerade alle auf die USA, weil eben vermutet wird,

dass wir hier so richtig Desinformationen erleben würden.

Und ich finde die Stimme erschreckend gut.

Das Einzige, was ich jetzt so finde,

ist, dass das nicht exakt zusammengepasst hat.

Dass der Ton mit ihrer sehr entschlossenen Stimme hat nicht ganz zur entspannteren Mimik gepasst.

Wenn ich mir jetzt vorstelle, in der Geschwindigkeit entwickelt sich das weiter,

dann könnte ich mir vorstellen, dass das glaubwürdig wäre.

Es reicht ja schon, Zweifel zu sehen.

Ach so, du meinst, wenn solche Videos im Umlauf sind,

dass Zweifel an echten Videos gesagt werden?

Europol hat jetzt gerade einen recht frischen Bericht dazu rausgegeben, wo die sich mal mit diesem ganzen Thema Deep Fake beschäftigt haben und wo die die Gefahren sehen.

Was ich ganz spannend finde, ist,

was wir gar nicht so auf dem Schirm haben,

ist eben, was bedeutet das für Ermittler?

Weil eigentlich so eine Zeugenaussage auf Band war immer ein guter Beweis

oder ein Geständnis, kennt man ja auch aus The Wire oder irgendwelchen Filmen.

Man überführt den Täter.

Inwiefern wird das vor Gericht dann noch Bestand haben, wenn es nur Eingeständnis gibt und das es als Audio verfügbar. Das heißt, ein Angeklagter oder seine Verteidigung könnte dann vor Gericht so eine Audioaufnahme als Beweis an Zweifel. Also, wofür führt das hin? Was macht das mit uns als Gesellschaft, wenn wir den Stimmen, die wir hören, nicht mehr trauen können? Ja, ich finde, das ist eine völlig neue Dimension und ich glaube auch, dass uns das die nächsten Jahre richtig beschäftigen wird. Also, gerade in diesem Feld desinformativen, also Beispiel Donald Trump und Scab, diese sehr bekannte Tonaufnahme, Grab and Buy the Pussies. Jetzt kam das wieder hoch im Rahmen eines Prozesses. Wenn der damals eben gesagt hätte, das war ein Deep Fake, dann wäre diese Tonaufnahme null und nicht dicht gewesen. Das war 2015, 2016, da war die Technologie noch nicht so weit. Das heißt, was passiert, wenn wir im Prinzip nichts mehr trauen können? Also, nehmen wir auch an, es ist vielleicht ein Hörfunk-Interview geführt worden mit einem Politiker oder einer Politikerin. Und die sagt dann aber, jedes Interview habe ich nie gegeben, das ist ein Deep Fake. Habe ich so gar nicht gesagt. Ja, es ist ein ganz gefährlicher Zustand und auch, glaube ich, kann jeder nachempfinden, ein ganz unangenehmes Gefühl, wenn man sich die Frage stellt, was kann ich überhaupt noch glauben? Man versucht natürlich, dem irgendwie herzuwerden. Das ist aber gar nicht so einfach, weil wir hier ein Stück bei einem Katzen-Maus-Spiel spielen. Es gibt natürlich Technologie, die versucht auch solche Deep Fakes zu entdecken oder zu erkennen, also quasi, ich sage mal, KI mit KI-Bekämpfen oder Technik mit Technik beikommen. Jetzt zum Beispiel bei dem Anbieter, bei dem wir jetzt den Test mit dir gemacht haben, Victoria, die machen so eine Art Wasserzeichen, bauen die ein in diese Stimmen. Also kannst du das Sample wieder hochladen und dann solltest du sehen können, aha, das ist quasi von uns gemacht worden. Jetzt schreiben die aber auch schon wieder auf der Webseite, wenn das halt dann noch mehrfach verändert oder wenn man sich Sex mal selber per WhatsApp schickt oder so. Dann verwäscht das wieder. Dann wird die Erkennungsrate nicht mehr sicher.

Und da helfen dann einfach Sachen, die jetzt gar nicht so viel in der Technik erst mal einen Moment zu tun haben, klar, wie halt der sofortige Anruf, wenn ich bei einem Anruf halt zweifel, ist der Konzern Chef das oder nicht. Und das geht natürlich nur, wenn ich mich davor nicht fürchte. Oder jetzt im Fall vom Enkeltrick, klar, sofort einfach die Enkelinn anrufen, da einen kurzen Draht haben. Oder auch aus Ermittlerperspektive, das stand auch in diesem Europolbericht drin, die sagen natürlich, klar, wir brauchen mehr Personal, Personal Schulen auf solche Fälle. Ich hatte mit dem LKA in Hannover Kontakt. Die haben sich jetzt mal angefangen dazu zu informieren. Also, dass man z.B. aus mehreren Perspektiven Vorfälle braucht, um ein Beweis zu führen. Also so, wie wir das Open Source Intelligence Journalismus auch machen. Also, wenn wir versuchen, Videos zu verifizieren, dann gehen wir ja auch so vor, dass wir sagen, okay, welche Anhaltspunkte gibt es noch außer dem Video, wo wir dann glauben können, ja, das ist wahrscheinlich von diesem Ort, zu diesem Zeitpunkt wirklich passiert. Also, dass man sich z.B. andere Videos noch beschafft, die zum gleichen Zeitpunkt aufgenommen worden sind, dass man auf Google Earth und Konkurrenzunternehmen sich den Ort noch mal genau anschaut oder persönlich, halt einfach quasi mehr Quellenprinzip für Videos und Ton, ne? Ja, genau. Und trotzdem glaube ich, dass dieses Glaubwürdigkeitsproblem auf uns geguckt, auf den Journalismus, das wird unser Thema werden. Und ich denke, eine Möglichkeit, mit dem wir damit umgehen können, und das versuchen wir natürlich auch, ist halt mit Transparenz, mit Offenheit, mit Zeigen, wie wir arbeiten. Mit, ja, mit auch Gesichterzeigen, die hinter den Produkten stehen. Also Prinzip, dass man sagt, okay, dem NDR als Marke ist, das fällt schon von so einem Interview nicht zu, zu trauen. Also wirklich so Stichwort, Marken und Glaubwürdigkeit. Wie ist es denn eigentlich mit Social Media? Das ist ja auch ein ganz, ganz großer Verteiler für Fake News und auch gefälschte Zitate, Videos und auch dann eben neuerdings Audio-Dateien aller Art. Wie ist das? Gibt es da Kontrollmechanismen? Es gibt ja z.B. auch eigene Abteilungen, z.B. bei Meta, die dann auch ne Warnung manchmal angeheftet haben an Posts,

dass da inhaltlich irgendwas nicht ganz stimmt,  
dass es Fake News sind.

Ist dem da so beizukommen oder wie weit sind die da?

Also wir haben da so ein bisschen im Moment die Hypothese,  
wir haben da so die besten Zeiten gesehen.

Also das war ganz groß in der Corona-Zeit, in der Pandemie,  
wo YouTube und also Google und auch Meta da auch sehr viel getan haben  
und auch so Warnhinweise gemacht haben  
und auch wahnsinnig viel Geld investiert haben  
in dieses ganze Thema Desinformation.

Und jetzt muss man so ein bisschen, sage ich mal,  
mit dem Adler-Auge hingucken.

Denn was wir gerade erleben, es hat ne Beginn in der Rezession,  
es werden sehr, sehr viele Leute entlassen.

Ich meine, bei Twitter, was ja im Prinzip fast täglich in den Nachrichten,  
dass dort Content-Moderatoren, also Leute,  
die genau solche falschen Videos aufdecken,  
die so was runternehmen,

da im ganz großen Stil sind da Leute entlassen worden,  
die sich genau um solche Sicherheitsfragen gekümmert haben.

Und man sieht es aber auch ganz still, heimlich, klamm und leise,  
bei den anderen großen Anbietern, also bei Meta und bei Google,  
dass auch die ihre Bemühungen, was das angeht,  
ein Stück weit zurückfahren.

Also das beste Beispiel jetzt,

ich hatte jetzt mit Meta ersten Interview kürzlich dazu,  
oder hatte ich dann auch mal gefragt, ja, wie ist das mit Lügen?

Dürfen sich Lügen auf eurer Plattform verbreiten?

Ja, gut, kommt drauf an. Nein, eigentlich darf man keine Lügen verbreiten.

Aber ja, was ist denn, wenn Politiker Lügen verbreiten?

Ja, ne, Politiker, die dürfen Lügen, ja.

Weil es sind halt Politiker und politische Kommunikationen,  
wollen wir nicht einschränken.

Da müssen wir hingucken als Journalistinnen und Journalisten,  
vor allem auch.

Okay, hier kam eine Sprachnachricht.

Hallo, das Passwort?

Nein.

Bei mir ist es einfach, dass ich mein eigenes Passwort eingebe,  
also meine E-Mail-Adresse.

Nein.

Ich muss sie jetzt eigentlich mal anrufen.

Okay, erlöse Sie.

Ich muss sie mal erlösen.

Hallo, es tut mir total leid,  
aber ich wollte das noch auflösen, weil wir dich gediebt faked haben.  
Wir haben dich gescannt.  
Ich hab's mir fast gedacht, weil das hat sich so komisch angehört,  
dass ich auch keinen Passwort verraten.  
Ja, da lagst du sehr richtig.  
Es war ein Scam, es tut mir sehr leid.  
Ich kann mir währenddessen wirklich auch schon nicht gut vor.  
Aber wir wollten es mal ausprobieren.  
Alles für 11 km.  
Ja, okay.  
Dann, in diesem Sinne, mach's gut, tschau.  
Ja.  
Wenn wir jetzt ein paar Folgen richtig gefüttert hätten  
in fünf Minuten wäre es vielleicht schon anders ausgegangen.  
Who knows?  
Also, wenn es um paar Millionen Euro geht  
oder um mehrere hunderttausend Euro,  
da kann ich mir schon vorstellen,  
da investieren Menschen.  
Wow. Okay.  
Svea, danke, dass du hier warst  
und uns davon erzählt hast  
und uns das alles gezeigt hast  
und meine Stimme gefaked hast.  
Sehr gerne, danke, dass du dich drauf eingelassen hast.  
Gerne doch. Bis dann.  
Tschüss.  
Und um zu demonstrieren,  
dass der 1 Euro-Klon mich noch nicht gleich arbeitslos macht,  
sprechen wir das Auto jetzt abwechselnd.  
Das war unsere Folge für heute über Audio Deepfakes.  
Mit der NDR-Journalistin Svea Eckert.  
Svea hat übrigens auch eine eigene Podcast-Reihe gemacht.  
She Likes Tech ist der Titel.  
Sehr empfehlenswert, findet ihr in der ARD-Audiothek.  
Den Link gibt's in den Shownotes.  
11 km findet ihr in der ARD-Audiothek  
und überall, wo es Podcasts gibt.  
Abonniert uns da gerne.  
Folgenautor ist Hans-Christoph Böhringer.  
Mitgearbeitet hat Hannes Kunz.  
Produktion Christoph von der Werf, Jonas Teichmann,  
Eva Erhardt und Christine Dreyer.

## [Transcript] 11KM: der tagesschau-Podcast / Deepfake: Bei Anruf Klon

Redaktionsleitung Lena Gürtler und Fumiko Lipp.  
Mein Name ist Victoria Michaelczak.  
Fast. Mein Name ist Victoria Michaelczak.  
11 km ist eine Produktion von WR24 und NDR-Info.  
Wir hören uns morgen wieder.  
Und wir, hoffentlich nicht.  
Tschüss.  
Übrigens, die News-WG vom BR  
hat letzte Woche selbst ausprobiert,  
ob man mit so einem Stimmklohn  
ein Familienmitglied reinlegen kann.  
Anders als bei unserer Folge  
geht's da nicht nur um eine Sprachnachricht,  
sondern um ein Telefonat,  
bei dem die Audioschnipsel vorher  
recht schnell mithilfe einer KI generiert wurden.  
Ob das klappt?  
Wir verlinken euch das Instagram-Real in den Shownotes.