



#MTM22

18.10.2022

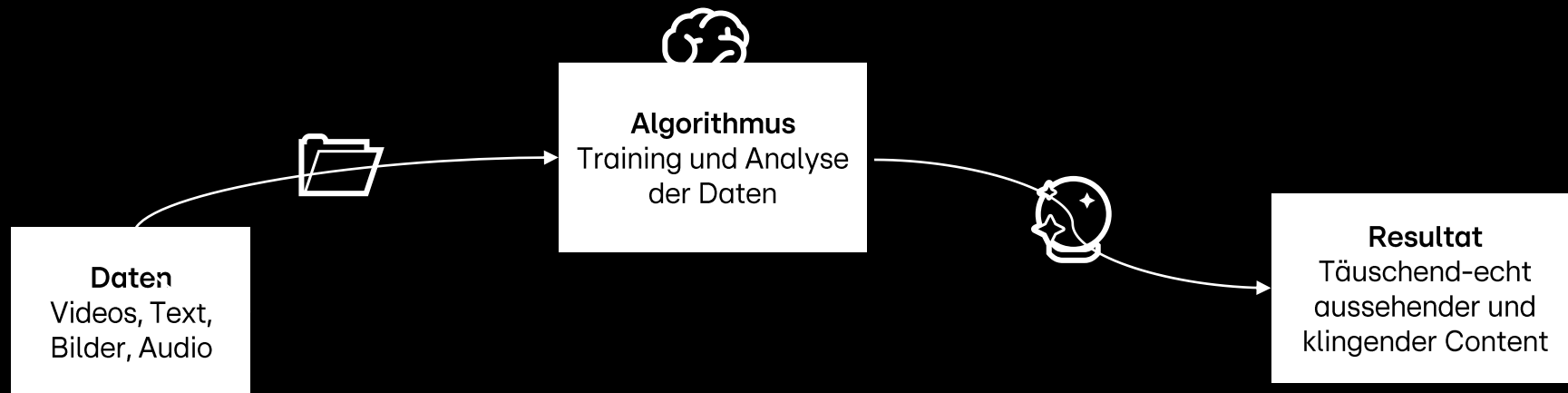
Clemens Siebler
Tatjana Anisimov

Synthetische Stimmen bei ntv

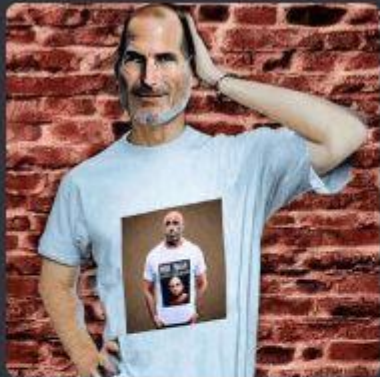
**Nicht nur Alexa: Sprachtechnologie als strategisches Instrument
für Reichweiten- und Produktentwicklung**

Was sind synthetische Medien?

Medien, die durch einen Algorithmus ganz oder teilweise erzeugt werden – insbesondere durch den Einsatz von künstlicher Intelligenz.



Welcome to [podcast.ai](#), a podcast that is entirely generated by artificial intelligence. Every week, we explore a new topic in depth, and listeners can [suggest topics or even guests and hosts](#) for future episodes. Whether you're a machine learning enthusiast, just want to hear your favorite topics covered in a new way or even just want to listen to voices from the past brought back to life, this is the podcast for you.



PODCAST.AI • EPISODE 1

Joe Rogan interviews Steve Jobs



1x



15:23 | 19:17

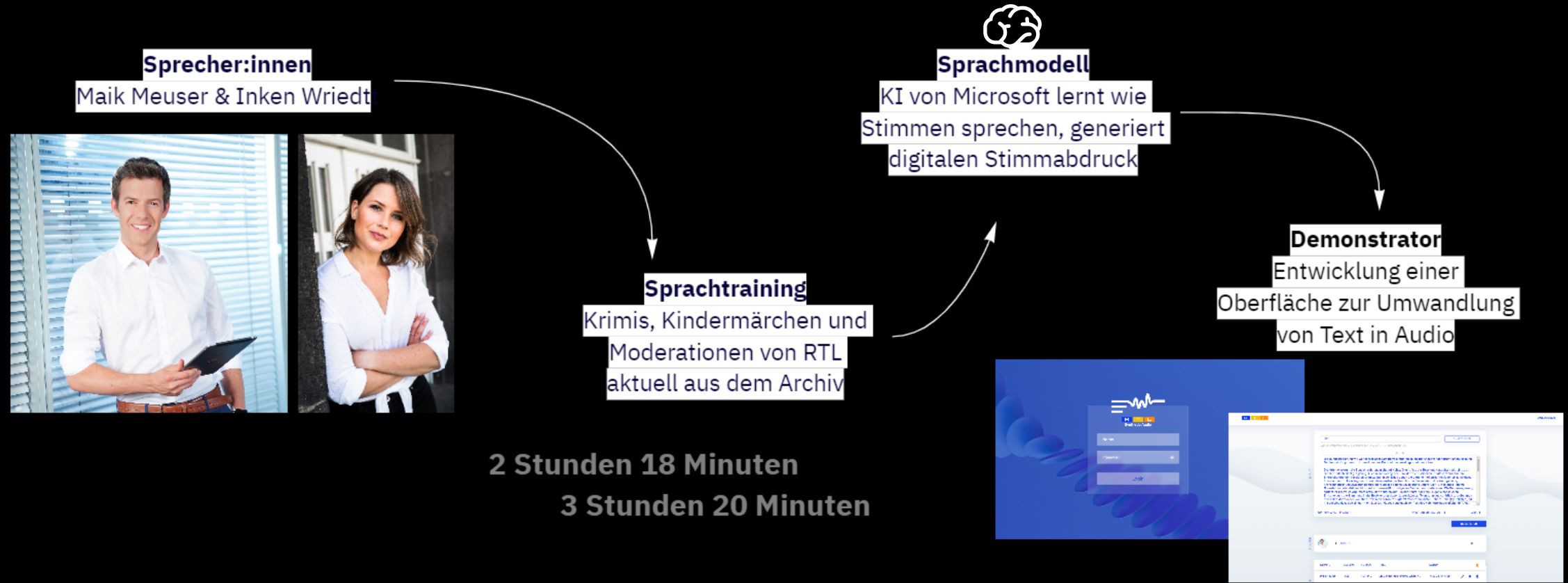
SUBSCRIBE

SHARE

MORE INFO



Transistor



Azure Neural Text-to-Speech

Very natural sounding voices

> 400 voices

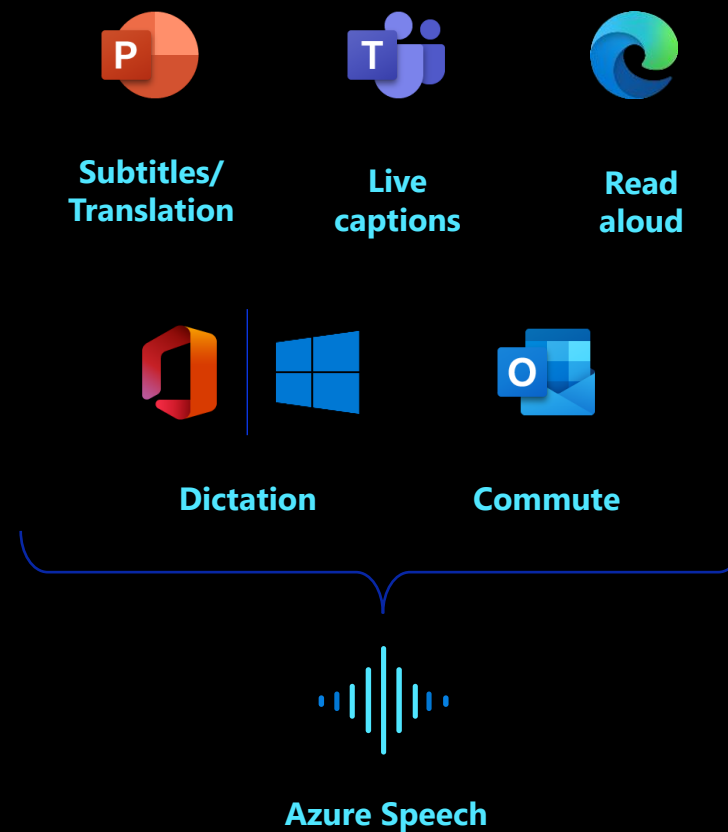
> 140 languages

Customizable with your own voice

Multi-lingual out of the box

No-brainer TCO for most use cases

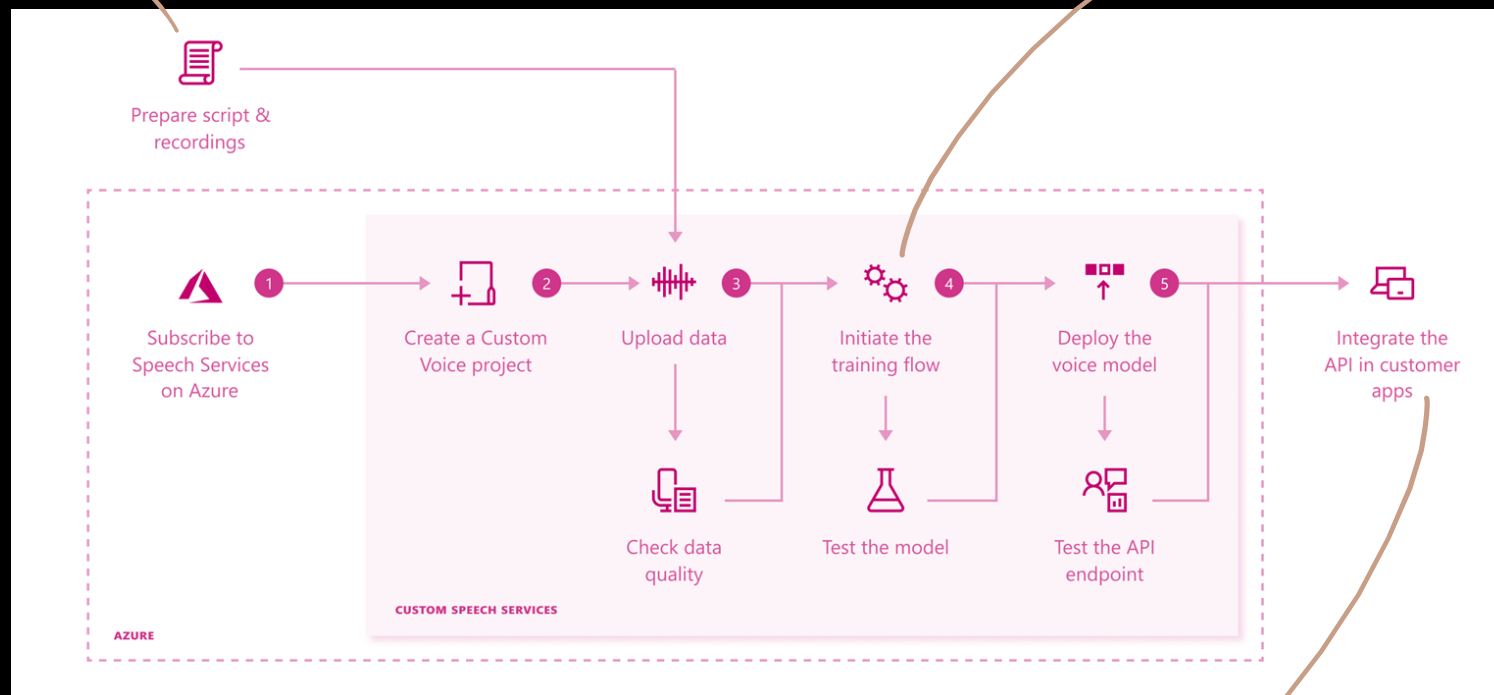
API or UI experience



Building a custom neural voice on Azure

Record speaker of choice
or re-use existing audio

Wait for training to finish (<24h)



Integrate voice into your
workflows

Synthetische Stimmen bei ntv

Unsere These

Nutzer:innen vertrauen unseren TV-Gesichtern und assoziieren ihre Stimmen auch auf auditiver Ebene mit RTL 

Use Cases



Nachrichten vorlesen



Videos vertonen
(Webvideos, Werbe- und
Audiospots)



Nachrichten im Radio zu
Randzeiten ansagen



Personalisierte &
lokalisierte Voice
Briefings

Hört ihr den
Unterschied?



**Die synthetische Stimme
unterscheidet sich kaum vom Original**



**Den Unterschied macht ein
transparenter Umgang und
verantwortungsvoller Einsatz
der Technologie**

Responsible use of synthetic voices

„Soft“ Solutions

- **Speaker consent**
- **Contractual agreement**
- **Code of Conduct**
No misinformation, manipulation, impersonation, spam, political
- **Approved Use Cases**
Education, news, marketing, entertainment, content creation, IVR/Voice bot, localization, accessibility, etc.
- **Abuse reporting**

Technical Solutions

- **Goal**
Identify synthetic voices “in the wild”
- **Watermarking**
With 
Without 
- **API to validate**
Check if audio came from Microsoft speech
Allows to distinguish custom voices from real speakers

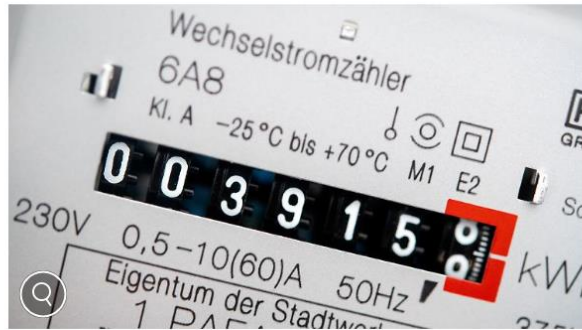
Startseite > Wirtschaft > Kommunale Prognose für 2023: Versorger erwarten 60 Prozent Strompreisanstieg

WIRTSCHAFT

Kommunale Prognose für 2023

Versorger erwarten 60 Prozent Strompreisanstieg

16.09.2022, 07:57 Uhr



Nachdem die Preise bisher um rund 10 Prozent gestiegen seien, wird für das kommende Jahr tendenziell eine Zunahme um die 60 Prozent erwartet. (Foto: Hauke-Christian Dittrich/dpa/Sym)

Audio player interface with a red border highlighting the playback controls and a note: "Diese Audioversion wurde mit Sprachrobotern unserer Moderatoren künstlich generiert. Wir freuen uns über Ihr Feedback zu diesem Angebot."

Die explodierenden Kosten für Energie bedrohen die Existenz der Stadtwerke. Daran ändern auch die jüngsten Gaspreissenkungen im Großhandel nichts. Der Verband kommunaler Unternehmen rechnet mit hohen Steigerungen der Strompreise für das kommende Jahr.

Deutschlands Stadtwerke rechnen nach Angaben des Verbandes VKU im kommenden Jahr mit erheblich höheren Strompreisen. Bisher seien diese zwar geringer gestiegen als die Gaspreise, etwa weil die EEG-Umlage weggefallen sei. Aber auch beim Strom seien weitere Steigerungen zu erwarten, sagte ein Sprecher des Verbandes kommunaler Unternehmen (VKU) der "Neuen Osnabrücker Zeitung".

Nachdem die Preise bisher um etwa 10 Prozent gestiegen seien, werde für das kommende Jahr tendenziell eine Zunahme um die 60 Prozent erwartet. Beim Gas müssten Endkunden der Stadtwerke derzeit häufig zwischen 30 und 60 Prozent mehr zahlen als vor dem Ukraine-Krieg. Es gebe aber auch Stadtwerke, die ihre Preise mehr als verdoppeln müssten, teilte er auch mit. Die jüngsten Gaspreissenkungen im Großhandel sind aus Sicht

NTV ZERTIFIKATE



NTV ZERTIFIKATE 13.09.2022 15:30 | 09:57 min
Aktienmärkte im Wandel
Die Top-Themen auf dem Deutschen Derivate Tag

VIDEOS WIRTSCHAFT



WIRTSCHAFT 18.09.2022 17:35 | 01:42 min
Vorreiter beim Stromsparen
So will Passau 20 Prozent weniger Energie verbrauchen



WIRTSCHAFT 18.09.2022 09:16 | 01:38 min
Digitalisierung auf dem Acker
GPS-Roboter "Dino" hackt Unkraut im Eiltempo



Startseite > ntv intern > Lassen Sie sich etwas vorlesen! Synthetische Stimmen bei ntv.de

NTV INTERN

Lassen Sie sich etwas vorlesen! Synthetische Stimmen bei ntv.de

Von Tatjana Anisimov und Samira Lazarovic 12.04.2022, 00:00 Uhr



Artikel anhören

Wie wäre es, wenn Sie sich ntv-Artikel von unseren Moderatoren vorlesen lassen könnten? In unserem Innovationsprojekt probieren wir genau das aus. Der Clou: Die Stimmen wurden automatisiert, das Vorlesen übernimmt eine Künstliche Intelligenz. Auch für Moderatoren ein neues Erlebnis.

Texte in lebensechte Sprache übertragen – darum geht es bei der Entwicklung sogenannter "text-to-speech"-Angebote. Aus einem geschriebenen Text wird eine gesprochene Stimme erzeugt, und zwar eine, die nicht einem Roboter gleicht, sondern von unseren menschlichen Stimmen nicht mehr zu unterscheiden ist.

RTL Deutschland hat für ein von der Landesanstalt für Medien NRW gefördertes Innovationsprojekt die Stimmen von RTL-Moderator Maik Meuser und Podcast-Host Inken Wriedt mithilfe einer selbstlernenden Speech-Technologie von Microsoft künstlich nachgebildet.

Das Ergebnis können Sie sich nun in einem ersten Test bei ausgewählten Artikeln hier auf www.ntv.de sowie in der ntv-App für iOS anhören.

Wie entsteht der digitale Stimmabdruck eines Menschen?

Für die Sprachsynthese werden zwei Dinge benötigt: Daten und Zeit. In unserem Innovationsprojekt haben wir von Maik Meuser und Inken Wriedt mehrere Stunden Audiomaterial verwendet und damit die Künstlichen Intelligenz (KI) von Microsoft zum Trainieren angefüttert.

Bei diesem "Training" lernt die KI wie ein Kleinkind: Sie bekommt das eingesprochene Material immer wieder vorgespielt und leitet daraus ab, wie ein Mensch spricht und seine Stimme zusammengesetzt ist. Im Vergleich zu einem Kleinkind braucht die KI dafür nur ein bis zwei Tage.



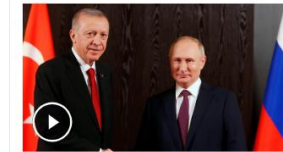
POLITIK 18.09.2022 09:46 | 01:55 min
"Alles in meiner Seele explodiert!"
Hinterbliebene sollen Leichen in Isjum identifizieren



POLITIK 18.09.2022 14:08 | 02:35 min
Abzug offenbart Kriegsgräuelt
ntv-Reporterin besichtigt russischen Folterraum



PANORAMA 18.09.2022 13:44 | 00:38 min
Schuldach stürzt im Unterricht ein
Schweres Erdbeben erschüttert Taiwan



POLITIK 18.09.2022 10:07 | 02:27 min
Türkei will Ost-Bündnis beitreten
Munz: "Erdogan braucht einfach außenpolitische Erfolge"



Unsere Takeaways

01

Custom Voice
macht den
Unterschied

02

Prominenz ist
sekundär

03

Qualität schlägt
Bekanntheit

04

Barrierefreiheit*
wird anerkannt

05

Fremdwörter
bringen uns an die
Grenze

06

Marathon kein
Sprint

Die Zukunft ist synthetisch und wir bewegen uns immer schneller darauf zu.

- Medien werden nicht nur physisch, sondern digital erzeugt werden
- Frage der Zeit, bis ganze Podcasts/Hörbücher in Hollywood Qualität nur mit einem Computer entstehen
- Bis 2030 werden 90 Prozent aller Inhalte synthetisch sein*, so wie wir es heute schon von Bilder kennen, die via Photoshop verändert wurden

Viel Potenzial...

- Menschen eine Stimme geben, die sie verloren haben
- Personalisierung im Gaming-Bereich (Charaktere)
- Menschen „wiederzubeleben“ oder auferstehen zu lassen (Andy Warhol Netflix Serie, Steve Jobs Interview)

...aber auch viel
zu diskutieren

**Wie soll diese synthetische
Zukunft aussehen, in der
wir leben wollen?**

DANKE!



Clemens Siebler
Senior AI Specialist
Microsoft

clemens.siebler@microsoft.com



Tatjana Anisimov
Product Innovation Manager
RTL Deutschland

tatjana.anisimov@rtl.de