

Emotionale KI in der Werbeforschung Beyond the Real Voice of the Customer

Malte Freksa
Sandra Vitt
Prof. Dr. Holger Lütters

GapFish GmbH
Mediengruppe RTL Deutschland GmbH
HTW Berlin



Wissen, was wirklich zählt



Hochschule für Technik
und Wirtschaft Berlin
University of Applied Sciences

Gemeinsame Arbeit zur Erforschung emotionaler KI Technologien in der Marktforschung

Das Projekt kombiniert neueste Technologien mit Anforderungen aus der Praxis der Werbeforschung

GAPFISH

Malte Freksa

**MEDIEN
GRUPPE
RTL**
DEUTSCHLAND

Sandra Vitt

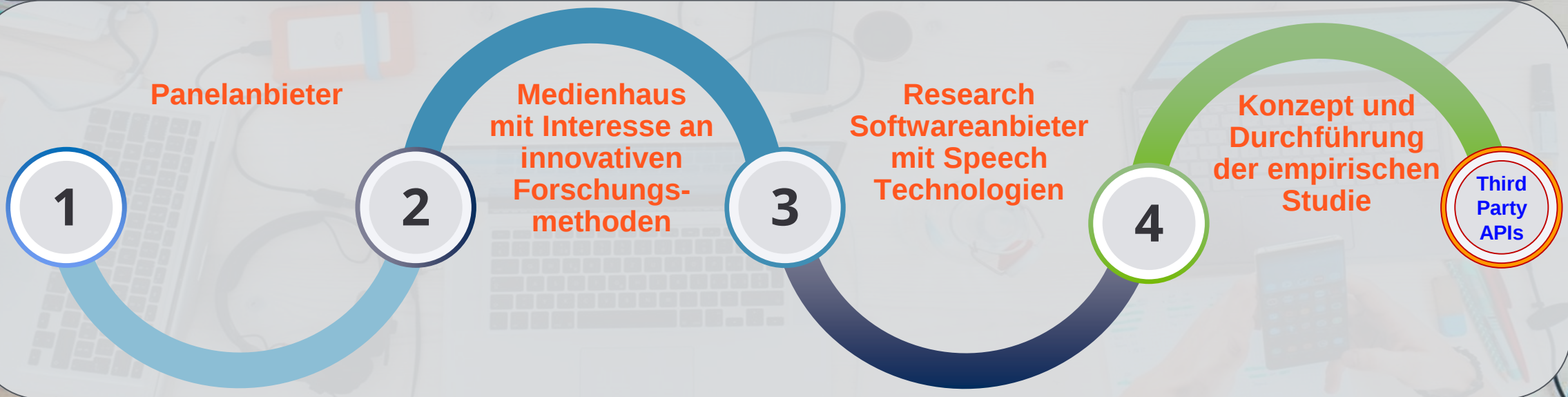
 pangea labs

Dmitrij Feller

htw

Hochschule für Technik
und Wirtschaft Berlin
University of Applied Sciences

Holger Lütters



Empirische Studie: Voice und Emotional AI in der Marktforschung



Welche Art von Kommunikation erzeugt welche emotionale Wirkung?

Moderne Werbeformate im Marktforschungstest mit modernsten Technologien

Marke Nutella 1 Spot

Marke Love2Say 4 Spots

2 Audio Spots

2 Instagram Posts



Können ≠ Wollen

Wir erklären heute das Verfahren und die Stolpersteine beim Einsatz Künstlicher Intelligenz im Interview der Marktforschung

Drücken Sie **START** und sagen nach Ablauf eines Countdowns bitte **"Das Pferd frisst keinen Gurkensalat"** ins Mikrofon.
Drücken Sie **STOP**, wenn Sie fertig sind.
Ihre Stimme wird danach in geschriebenen Text umgewandelt und direkt dargestellt.
Bitte beantworten Sie danach die untere Frage, wie gut die Umwandlung auf Ihrem Gerät funktioniert hat.



Wurde Ihr gesprochener Text erfolgreich umgewandelt?

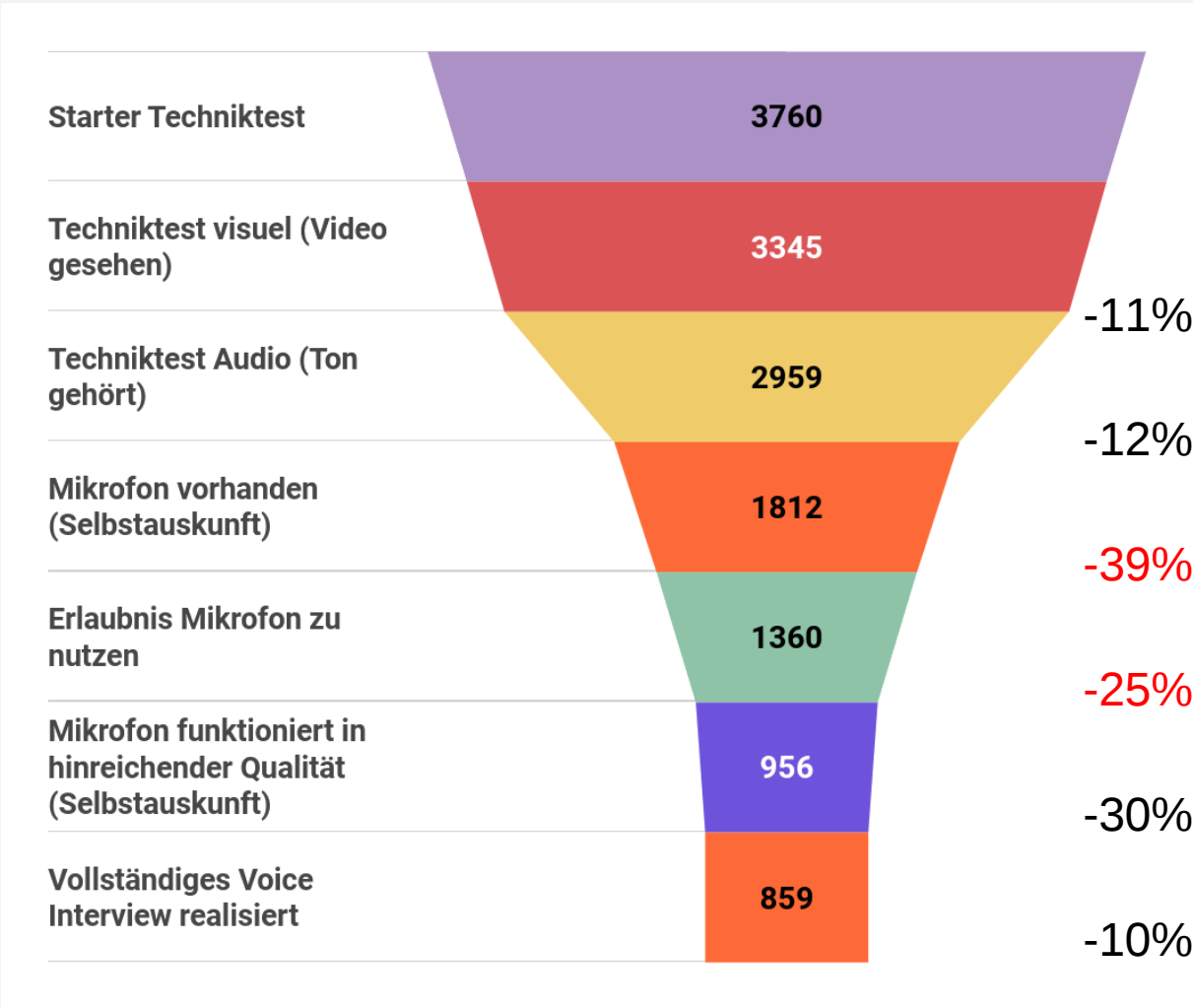
	Ja, die Umwandlung in Text funktioniert einwandfrei
	Die Umwandlung enthält leichte Fehler
	Die Umwandlung ist sehr schlecht
	Die Umwandlung funktioniert überhaupt nicht

Feldbericht

- Online-repräsentative Stichprobenziehung im EntscheiderClub Panel von GapFish
- Feldzeit: 2.-18. September 2019
- n=859 completes
- Durchschnittliche Zeit (Median): 14 Minuten 15 Sekunden

Können ≠ Wollen

Technik und mangelnde Teilnahmebereitschaft sind die Hauptherausforderungen bei der Umsetzung von Voice Befragungen



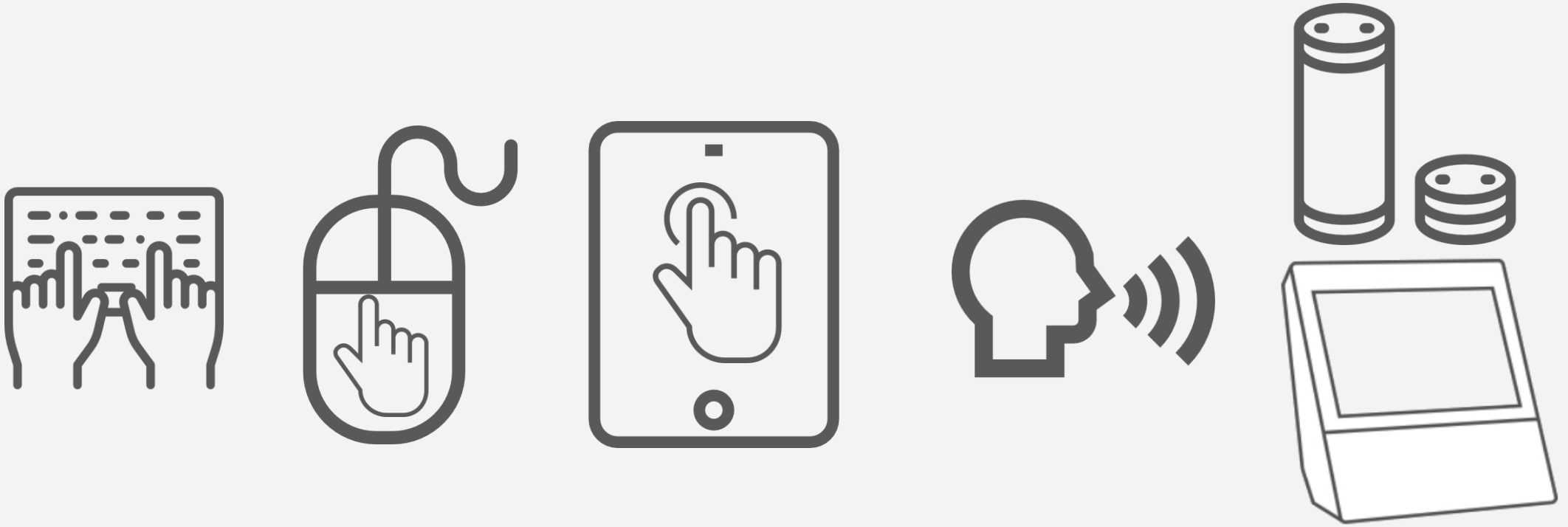
Wir sprechen schon immer mit Geräten

Verwendung gesprochener Sprache in Konsumenteninterviews



Wir leben in Zeiten der Veränderung von User Interfaces

Speech wird gerade zu einem (weiteren) Standard in der digitalen Interaktion



Trivialer Gedanke 2017: Warum also nicht Sprache in digitalen Interviews der Marktforschung einsetzen?

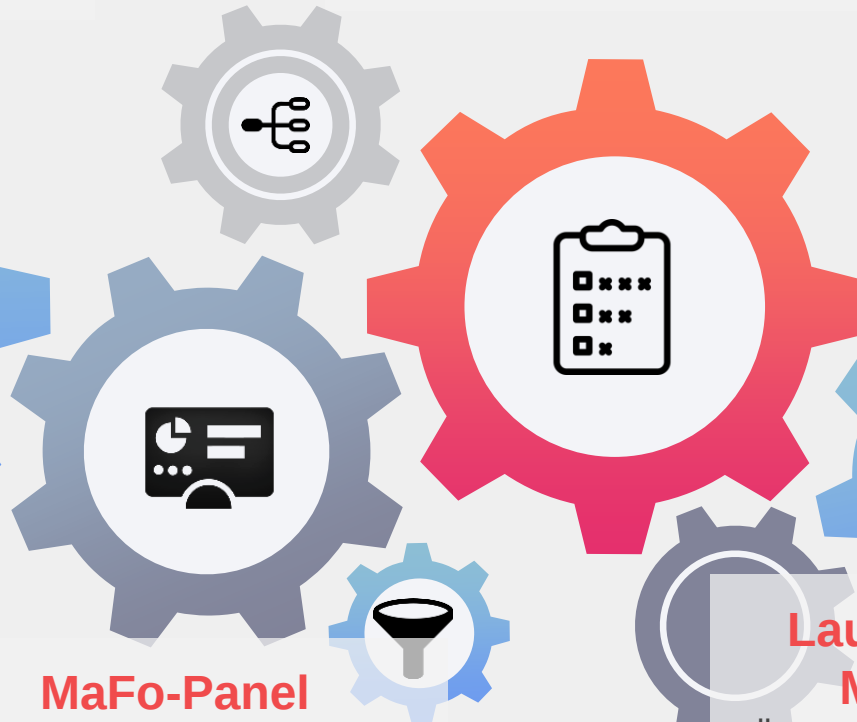
Verwendung von Speech2Text UND Emotion API im Marktforschungsprojekt mit Panelisten

Idee: BYOD erlaubt die Nutzung der bestehenden Infrastruktur der Marktforschung

BYOD
Bring your own device



Standard Fragebogen
questfox mit zusätzlichen
Voice Fragetypen



Speech2Text (STT)
Aufnahme von Audio Input mit
automatisierter Transkription

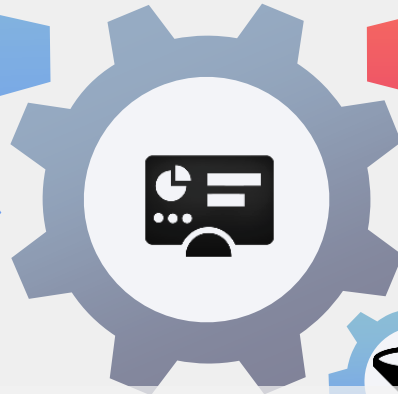


Emotion AI
API zur Überprüfung der
Emotionen aus Audio



MaFo-Panel
Infrastruktur

mit etablierten Prozeduren für
Sampling mit Filter, Quoten etc.



**Lautsprecher- und
Mikrofon-Test**

Überprüfung Bewegtbild / Audio



Umsetzung der Befragung ohne Programmierkenntnisse möglich



TTS
Text2Speech

Audio
Recording

Talk2Me
Speech2Text

Background
Speech2Text

Transcription
Quality Score

Video
Insights

Emotional
AI Voice
Analysis



questfox

8 unterstützte Sprachen (41 Varianten wie z.B. kanadisches Französisch etc.)

Englisch (13 Varianten), Deutsch, Niederländisch, Französisch (2), Portugisisch (2), Mandarin (2), Koreanisch, Spanisch (19)

Die Sprachaufnahme wird mit verschiedenen API Services weiterverarbeitet

Die Integration von API mit künstlicher Intelligenz wird live in den Befragungsprozess integriert und kann noch während der Befragung genutzt werden



Interview

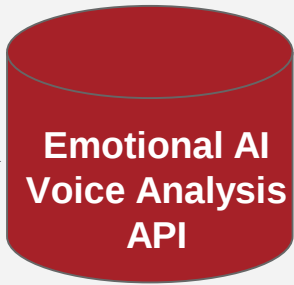
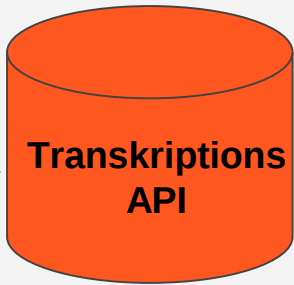
Aufzeichnung
Voice

API Integration
KI Technologien

Live Integration API
Ergebnisse in Fragebogen

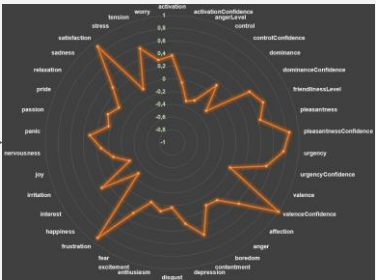
Nutzung der klassischen Online Research Infrastruktur mit Sampling inkl. Quotierung und Filterung

questfox Fragebogen mit Voice Elementen
Filterung der Menschen, die technisch **nicht** in der Lage sind per Voice zu antworten



- Transkribierter Text
- Antwortzeit
- Transkriptions Qualitätsscore

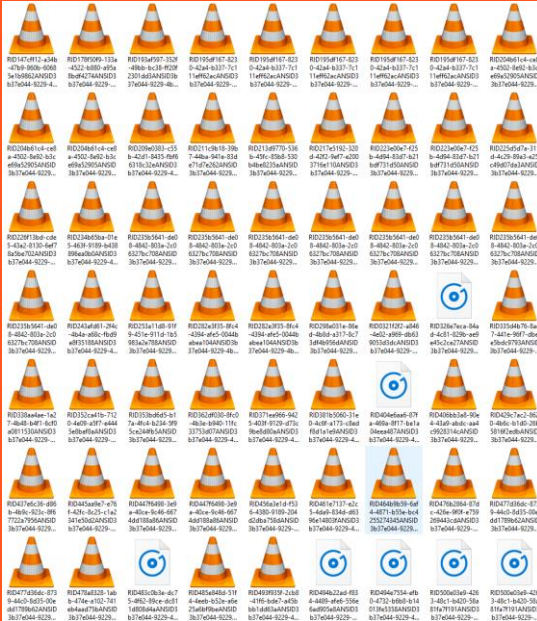
Ermittlung emotionaler Facetten aus der Stimme



Die neue Technologie erweitert die möglichen Auswertungen

Die Monologanalyse ergibt einen reichhaltigen Datenfundus auf verschiedenen Ebenen

Audio-Datei von jeder Antwort (optional)



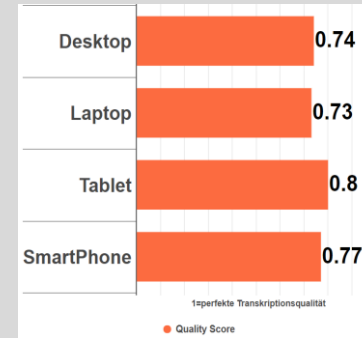
Automatisch transkribierte Texte

das pferd frisst mein gurkanalat wer frisst keinen gurkanalat
das pferd frisst keinen kartoffelalat wer frisst einen gurkanalat
erste bank open okay das pferd frisst keinen gurkanalat gurkanalatstschüss
keinen gurkanalatdas pferd frisst keinen gurkanalat wer ist nico
das werkzeug auto salat lalala das pferd das pferd frisst keinen gurkanalat geht nicht
das pferd frisstdas pferd frisst keinen gurkanalat das herz frisst keinen gurkanalat
jetzt frisst keinen gurkanalat nein das pferd frisst keinen kartoffelalat 123
das pferd frisst keinen gurkanalat das pferd frisst keinen gurkanalat
das pferd frisst keinen gurkanalatdas pferd frisst keinen gurkanalat ja das pferd
das pferd ist keinen gurkanalat dass das pferd frisst keinen gurkanalat
hallo das gut pferd frisst keinen gurkanalat jo super 0
das pferd frisst kein gurkanalat das pferd frisst keinen gurkanalat.
das pferd frisst das pferd frisst keinen gurkanalat verstärktes kann gurkanalat
pferd ist keinen gurkanalatdas pferd ist keinen gurkanalat hallo wie geht es dir
das pferd frisst keinen tosdas hört frost kein gurkanalat was sagt ein gurkanalat

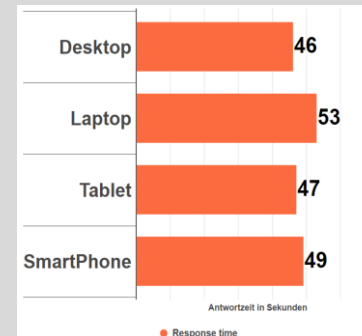


Paradata

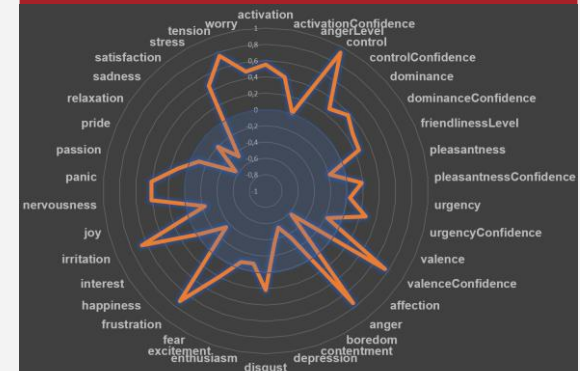
Transcription Quality Score



Antwortzeit



API Auswertung Emotionaler Facetten



Zu jeder Art von Stimulus kann der Proband seine verbale Meinung äußern

Beispiel: Antwort nach Konfrontation mit Radiospot der Marke Nutella



Audio-Datei von jeder Antwort (optional)



Automatisch transkribierter Text

also für Radiospots habe ich ja auch so gar keine Vorliebe ich **ge**höre zwar gespannt zu und ich lasse das auch auf mich wirken aber eine bildliche Vorstellung von der Werbung passt mir dann schon eher ja die Werbung an sich also **den den** Radiospot an sich selber jetzt von Nutella fand ich soweit ganz informativ und sehr gut gestaltet ja mehr muss man dazu nicht sagen auch wieder so kurz und knackig wunderbar

Transkription korrekt aber

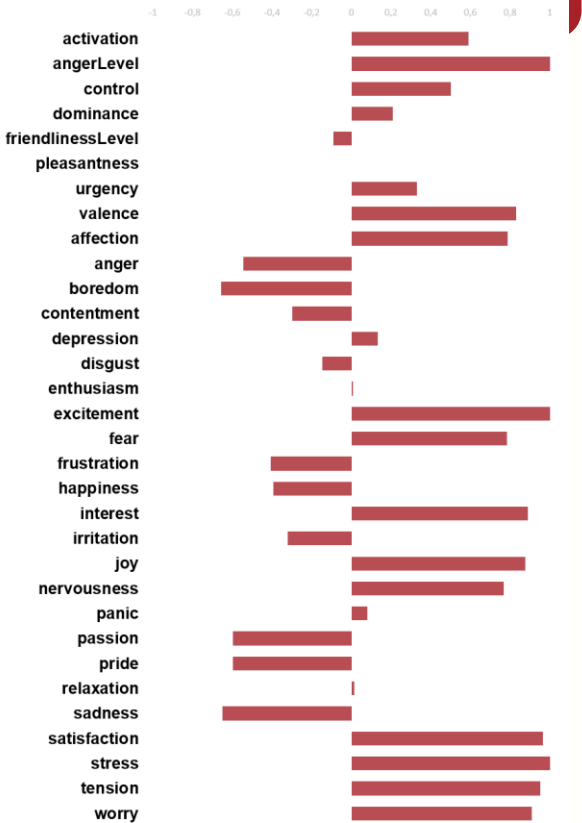
Paradata

Transcription Quality Score 0,92 bei Testfrage
0,81 bei der konkreten Frage

Antwortzeit 32,79 Sekunden

Stichprobe im Vergleich:
Kürzeste Antwort 1,9 Sekunden
Längste Antwort 48 Sekunden
Median 6,18 Sekunden

API Auswertung Emotionaler Facetten



Ein sehr konkretes Problem

Werbewirkungsforschung RTL

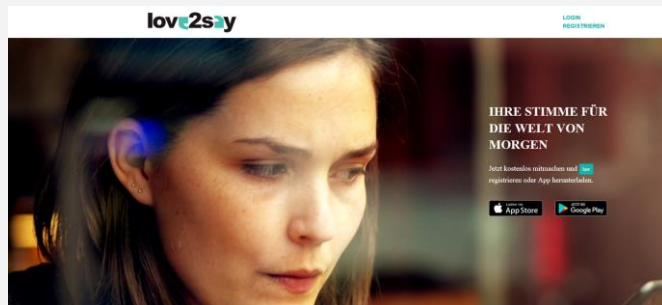
3



NEU seit April 2019: Unser neues Panel „Love2say“

 Stationär und Mobil

Aktuell 10.000 Teilnehmer



Verbesserte User-Experience:
Glücksrad und Erfolgserlebnis
am Ende jeder Befragung



Kernfrage: Auf welchen Spot fokussieren wir uns?

Rekrutierung für Love2say läuft im Fernsehen und digital mit verschiedenen Spots

Amiaz Habtu VOX Moderator



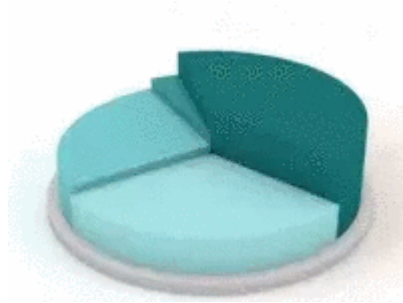
Wolfram Kons RTL Moderator



Klassische Befragungsdimensionen

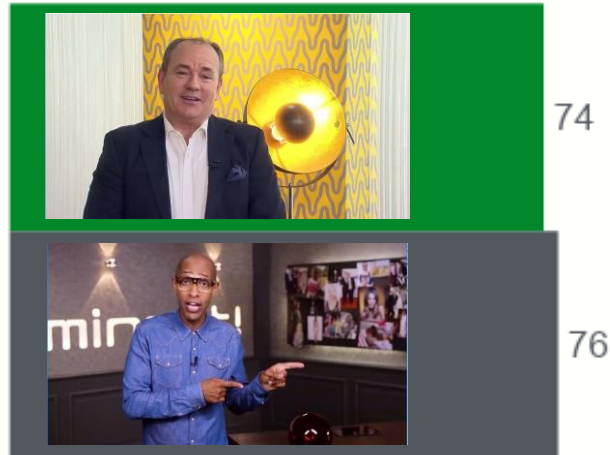
Auswertung TV Spot in klassischen Dimensionen

4



Wirkungstest Teil 1 - Sympathie

Der Moderator ist sympathisch



Amiaz Habtu

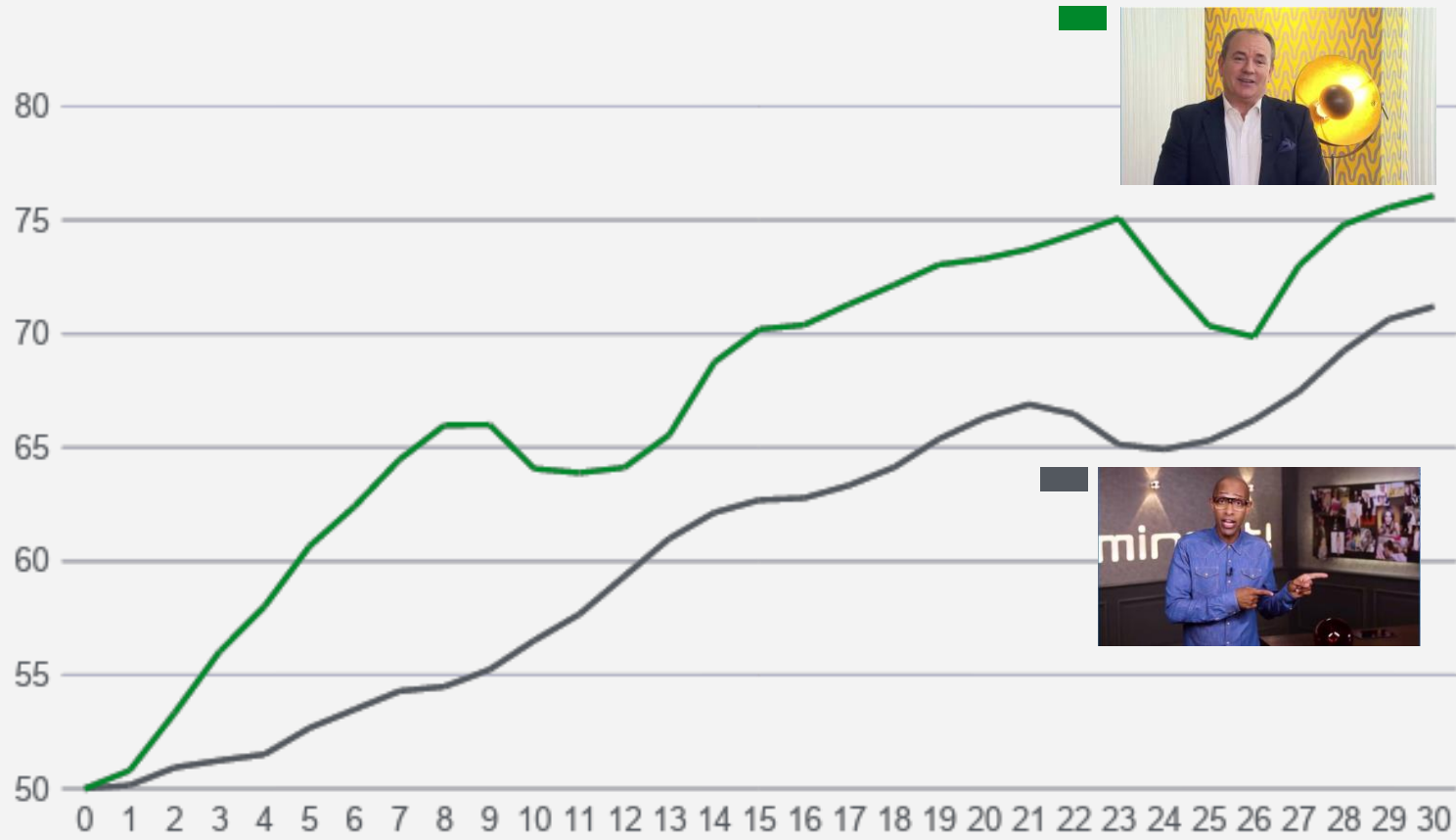
Wolfram Kons

1 : 0

Quelle: Mediengruppe RTL Deutschland, n=1.436, Angaben in %, Top-2-Box

Frage: Bitte bewerten Sie den Spot anhand der verschiedenen Aussagen. Bitte verwenden Sie dafür eine Skala von 1=stimme voll und ganz zu bis 4=stimme überhaupt nicht zu.

Wirkungstest Teil 2 – Spotbewertung im Sekundenverlauf



Amiaz Habtu Wolfram Kons

1 : 1

Quelle: Mediengruppe RTL Deutschland, n=1.436, Angaben in %, Top-2-Box
Frage: Bitte bewerten Sie das Video, inwieweit es Ihnen gefällt oder nicht, was Sie da sehen, indem Sie den Regler permanent hin und her bewegen.
Versuchen Sie, jede Änderung Ihres Gefallens mit dem Regler zu erfassen.

Wirkungstest Teil 3 – quantitative Spotbewertung



- ist originell
- ist vertrauenswürdig
- ist sympathisch
- ist außergewöhnlich
- ist verständlich
- macht Lust auf Teilnahme
- hat Charakter/ Persönlichkeit
- bringt mich zum Schmunzeln
- ist kreativ gemacht
- vermittelt alle wichtigen Informationen
- bietet etwas Neues
- hebt sich von anderer Werbung ab
- passt zum Thema
- passt zur Marke love2say
- stellt die Marke love2say deutlich heraus
- macht auf die vielfältigen Möglichkeiten von love2say aufmerksam

Amiaz Habtu	Wolfram Kons
6,08	5,99
6,56	6,42
6,82	7,02
5,78	5,87
7,76	7,5
7,02	6,75
6,22	6,33
4,47	4,69
6,04	6,26
6,72	6,74
6,54	6,47
6,1	5,95
6,93	7,01
6,78	6,55
7,34	7,26
6,10	6,13

Amiaz Habtu Wolfram Kons

2 : 2

n=868
Frage: Bitte bewerten Sie den vorhin gesehenen Spot anhand der folgenden Aussagen.
Skala: 0=trifft überhaupt nicht zu bis 10=trifft voll und ganz zu

Auf der Suche nach der Entscheidung

Gibt es weitere Indikatoren bei der Beantwortung der Frage nach dem besten Spot?



Emotional AI

Beyond The Real Voice of the Customer

5

Data has a better idea

Zufällige Zuteilung zu einer der 4 Gruppen mit Speech und Emotionsmessung aus der Stimme

Zwischen 6 und 8 Sekunden liegt die durchschnittliche Antwortzeit beim Voice-Fragetyp

			Antwortzeit				
	n	Transcription Confidence Score	Mean	Std Dev.	Median	Min	Max
	126	0,84	8,1	6,0	6,4	2,1	37,1
	132	0,84	9,2	6,4	7,1	2,1	33,1
	157	0,82	8,8	6,1	6,9	2,1	32,3
	139	0,83	9,9	6,9	8,1	1,9	37,0

Zu jeder Art von Stimulus kann der Proband seine verbale Meinung äußern
Speech ist nicht nur für junge Leute!
Beispiel: Spot 30 Sekunden mit Testimonial Amiaz Habtu



Audio-Datei der Antwort (optional)

Automatisch transkribierter Text

Paradata

API Auswertung Emotionaler Facetten
n=1

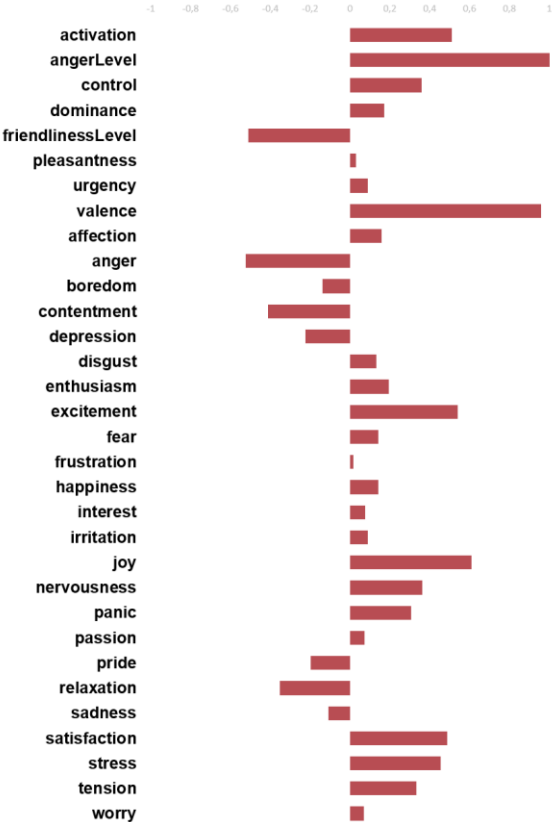


ja insgesamt hat er mir gefallen aber auch wieder alles nur sehr junge leute
meine generation wird immer außer acht gelassen
obwohl ich auch sehr interessiert bin an sowas und vielleicht sogar ein bisschen mehr dazu sagen würde als junge leute

Transkriptionsfehler

Transcription Quality Score
0,92 bei Testfrage
0,73 bei der konkreten Frage

Antwortzeit bei der Frage
25,02 Sekunden



Ein neues Datenset an emotionalen Zuständen steht zur Auswertung zur Verfügung

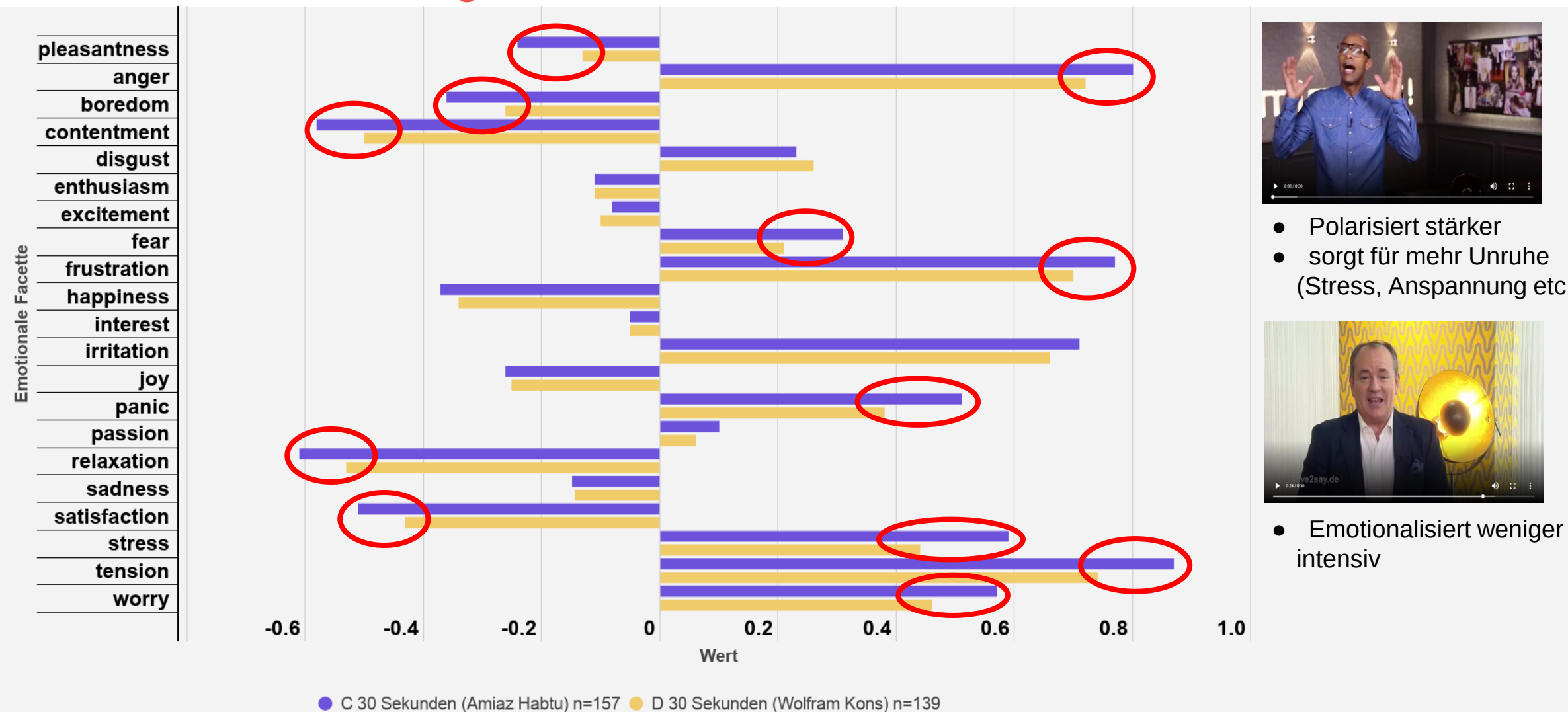
Die längeren Spots hinterlassen stärker ausgeprägte emotionale Statements



Kardinalfrage: Welche emotionalen Facetten sollten für Werbewirkung herangezogen werden?

Die Emotionsanalyse ergibt Mehrwerte in Bezug auf die Bewertung der Spots

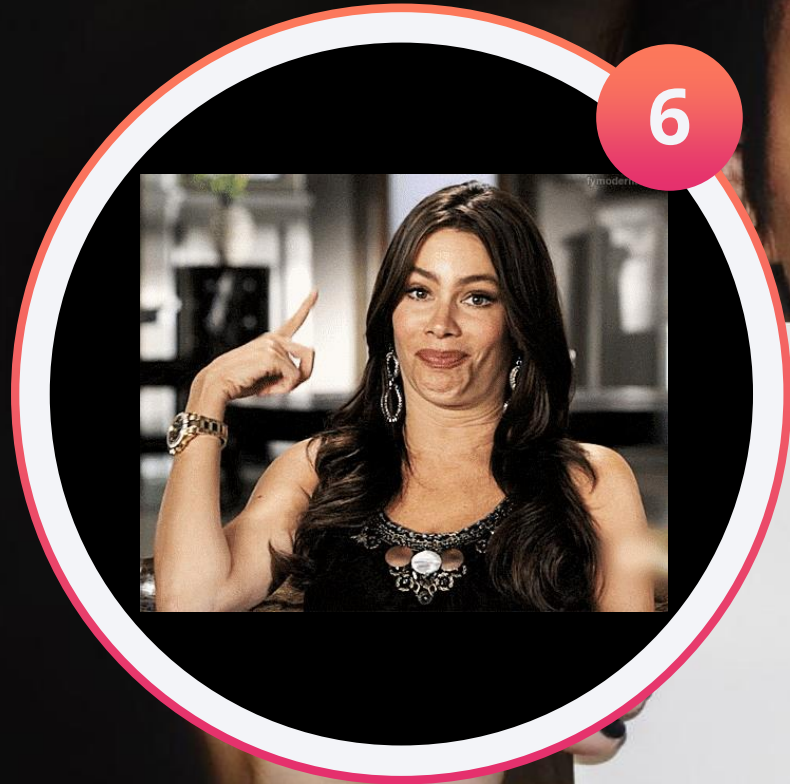
Die konkrete Entscheidung kann dadurch unterstützt werden



- Polarisiert stärker
- sorgt für mehr Unruhe (Stress, Anspannung etc.)



- Emotionalisiert weniger intensiv



Emotionale KI scheint enormes Potential für die Werbewirkungsforschung und andere Gebiete der Marktforschung zu haben. Eine kritische Validierung ist aber erforderlich.

Technische Fragestellungen beherrschbar

- Hoher technischer Drop-Out bei Audio-Interaktionen
- Nicht antworten wollen > nicht antworten können
- Einfluss der externen Settings MaFo auf die emotionale Situation besser verstehen bzw. stärker kontrollieren (im Sinne von Laborsettings)

Direkte Einsatzmöglichkeiten in der MaFo

Sofort umsetzbare Ideen:

- Brand Evaluation
- Customer/Employee Satisfaction
- Research Quality Control (Fraud detection)
- Multi-Language Tests
- Pre-Testing in kurzen Zyklen (Agile Management)
- ...

Nach der technischen Evaluierung steht die Validierung in verschiedenen Szenarien an. Das disruptive Potential ist hoch

Walking into the Next Generation of Research Designs



- **Die KI fügt einen weiteren Baustein der Analysemöglichkeiten hinzu**
- Qual und Quant könnten verschmelzen durch digitale Technologien
- Wir Marktforscher müssen zunächst neu lernen Fragen zu stellen
 - Ist unser gewählter Ansatz überhaupt richtig?
 - Welche Items sind adäquat?
 - Ist die klassische Statistik hier hilfreich, oder brauchen wir neue Metriken?
- Emo KPIs? Kampagnenspezifische Ziele der Emotion erforderlich?
- Hypothesen zur Emotion vorher entwickeln, bevor Studien im Data Mining enden
- Mensch noch einige Zeit erforderlich zur Analyse. Aber wie lange noch?

Wir danken den Unterstützern dieser Studie

Fragen, Diskussion und Vorführungen

Halle 2, Stand 205 bei GapFish

Projektdurchführung

Malte Friedrich-Frekša - GapFish GmbH
Sandra Vitt - Mediengruppe RTL Deutschland
Prof. Dr. Holger Lütters – HTW Berlin

Die Befragung mit Voice erfolgte mit www.questfox.com
von pangea labs unter Nutzung der in questfox
integrierten Emotional API von audEERING



https://www.love2say.de/users/sign_up