

Audio-KI im Gesundheitswesen

Erfahrungen eines KI-ScaleUps

Dagmar M. Schuller

WT23, 28.07.2023



audEERING's Voice AI has been recognized and featured by:

Forbes

 NATIONAL
GEOGRAPHIC



REUTERS

msn

ZDF

ORF **SRF**

ARD

Süddeutsche Zeitung

GEO

Frankfurter Allgemeine

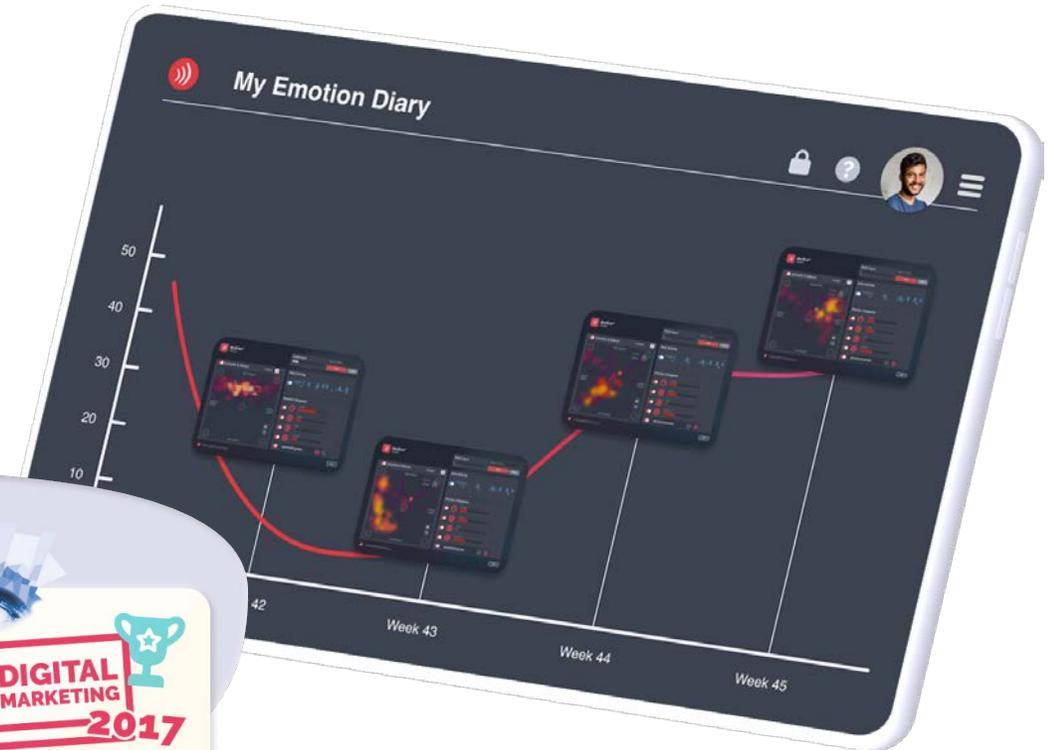


Erkennen und Verstehen von Zuständen und Merkmalen aus der Stimme

- in Echtzeit
- objektiv
- wissenschaftlich validiert
- klinisch validiert*

mit unserer **multi-Award-winning Voice AI** Technologie

* Auswahl von Markern im Bereich Mental Disorder, MS, etc.





No 1 Emotionserkennungsmodell aus der Stimme

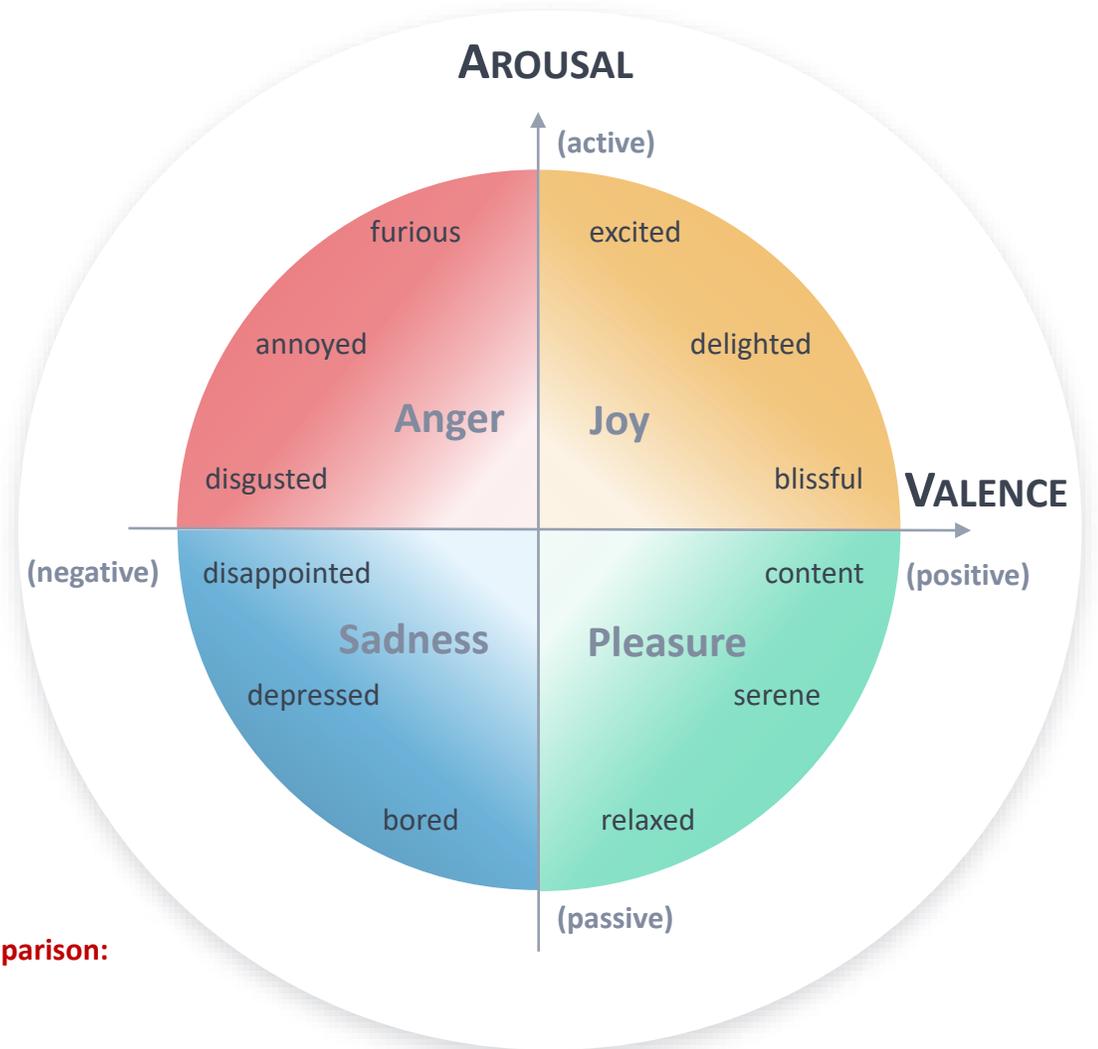
Tone-based Erkennung

unabhängig von Inhalt und Sprache

Aktuelle Performance gemessen als
Concordance Correlation Coefficient *1

	audEERING	Others
Arousal	0.75	0.6 – 0.7
Valence	0.64	0.2 – 0.5

*1 range 0 to 1: 1 = perfect correlation/ no error between reference and prediction
0 = completely uncorrelated /high error between prediction and reference; **HUMAN comparison:**
0.7





Cutting-Edge Technologie

- State-of-the-Art KI mit tiefen neuronalen Netzen und unsupervised / semi-supervised Learning
- Nahezu 7.000 interpretierbare Stimmfeatures messbar
- Zuverlässige KI, trainiert auf Millionen Stunden von Audio, ausgewogener Datensatz
- Light-weight Code, Fokus auf geringen Ressourcenverbrauch

```
38 // Initialize devAice VAD module
39 DeviceVADConfig vadConfig = DEVICE_VAD_CONFIG_DEFAULT;
40 vadConfig.min_segment_length = 0.45;
41 vadConfig.segment_start_delay = 0.1;
42 vadConfig.segment_end_delay = 0.1;
43
44 // Add VAD module
45 std::vector<DeviceModule> modules;
46 modules.push_back({ DEVICE_MODULE_VAD, &vadConfig });
47
48 // Initialize session
49 DeviceConfig config = DEVICE_CONFIG_DEFAULT;
50 config.resource_root_path = resourceRootPath;
51 config.num_modules = modules.size();
52 config.modules = modules.data();
53 DeviceResult result = devaice_init(session, &config);
54 checkResult(result, session);
55
56 // Setup audio file input
57 DeviceFileSettings fileSettings = DEVICE_FILE_SETTINGS_DEFAULT;
58 fileSettings.file_format = DEVICE_FILE_FORMAT_WAV;
59 result = devaice_stream_set_file(session, DEVICE_SID_AUDIO_INPUT, wa
60 checkResult(result, session);
61
62 // Setup VAD segment output stream in callback mode
63 DeviceStream *vadSegmentsStream;
64 result = devaice_stream_open(session, DEVICE_SID_VAD_RESULTS_SEGMENTS
65 checkResult(result, session);
66
67 // Run devAice processing
68 result = devaice_run(session);
69 checkResult(result, session);
70
71 // Get output size
72 size_t count;
73 result = devaice_stream_get_data(vadSegmentsStream, nullptr, &count);
74 checkResult(result, session);
75
76 // Get VAD output
77 std::vector<DeviceVADSegment> vadSegments;
78 result = devaice_stream
79 checkResult(r
80 for
```



Case Study: Schizophrenia Biomarkers

Ziel: Prognose von **Grad der Schizophrenie** und **Entwicklung / Relapse**

Daten: Vom Kunden zur Verfügung gestellt
aufgenommen auf Basis von 8kHz Telefoninterviews

Patienten mehr als 1 Jahr beobachtet
durch monatliche Stimmanalyse und klinische Diagnose

Ergebnis: Clinical Global Impression Scale (CGI-S)

87% Erkennung bei der Vorhersage von Intra-Patient CGI-S
Scores nach Training auf Basis von vorherigen
Patientendaten

Long-Term-
Study

Prediction of Schizophrenia severity
Confusion matrix: CGI-S (as a 2 class problem)

True label	Predicted label	
	low	high
low	1030	82
high	64	210

Performance of different **classification approaches** for different clinical scales

Table 1: Segment-level UAR results on the different targets using speaker-dependent splits for XGBoost with ComParE2016 features and CNN10.

Target	XGBoost	CNN10
CGI-S	.77	.87
RASP	.67	.73
RASP-A	.66	.72
SRF-PANNS	.70	.75



PROJECT

Mental Wellbeing at Sea

- **Crewmitglieder eines Öltankers**

Aktive Aufnahmen: Daten über AI SoundLab

Passive Aufnahmen: Reisedaten Recorder(black box on ship's bridge)

- **Aktive Daten : Sprachtests**

- Freie Daten, sustained utterance, emotion acting

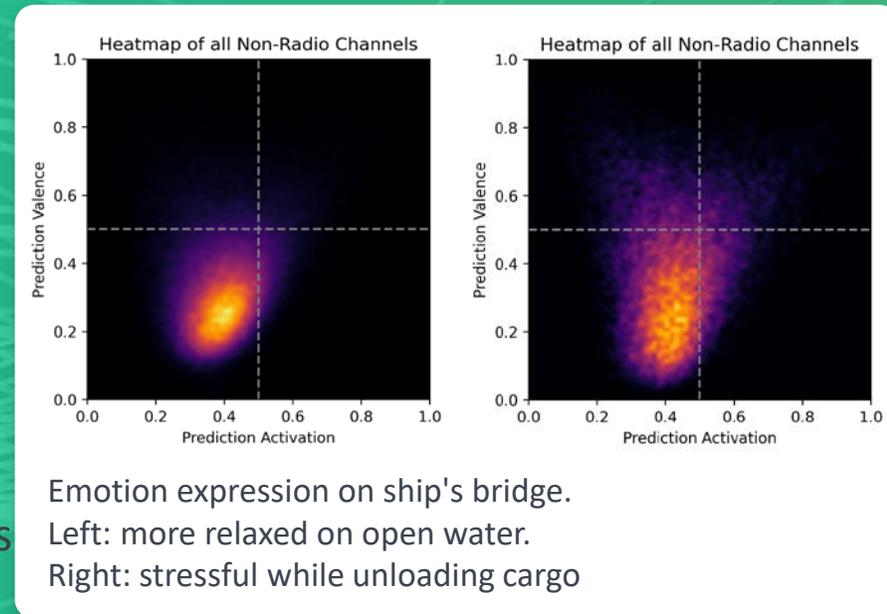
- **Passive Daten**

- Stressmomente: higher spread and variance in emotion express



AI Results Stress Level Erkennung:
approx. 70%+ accuracy* vollautomatisch; = human level

*Pearson Correlation Coefficient

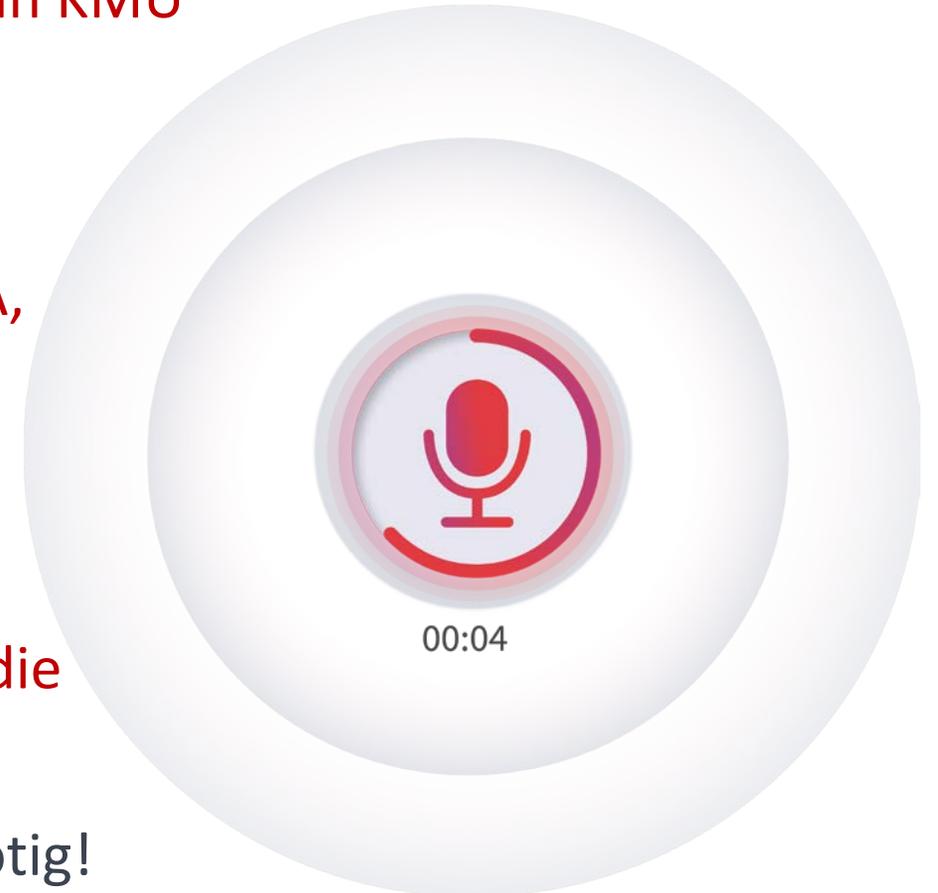




Herausforderungen...



- Überregulierung: Daten, AI-Act, Anforderungen an KMU sehr hoch
- Risiko-averse VC Struktur
- StartUp-Phase – kein Problem, Scale-Up (Series A, Series B) schwieriger als in anderen Märkten
- Öffentliche Fördermittel sind nicht wachstumsfreundlich in der Umsetzung
- “Geht das überhaupt?” – vs “Wann erobert Du die Welt?”
- Momentum im KI-Bereich: starkes Investment nötig!





...und wie man sie lösen könnte...

- Regulierung anpassen an die Realität: mehr Pragmatismus und Unterstützung der lokalen Player
- Mehr Offenheit und Unterstützung für Investments im Bereich EUR 20 Mio – EUR 50 Mio Series A oder Series B (bspw durch steuerliche Incentives, eigene Programme)
- **Selber nutzen und darüber sprechen!**
- Stärkerer Fokus auf Wachstumsphase und Technologietransfer bei Förderungen
- **Prozessverbesserungen im öffentlichen Bereich**



Vielen Dank!

