

Speech-Based Cardiorespiratory Health Monitoring with VOICE-BIOME: a scalable voice biomarker platform (Practical Contribution)

Prashanth Pombala¹, Lars Nippet¹, Julia Hoxha¹ and Sami O. Simons²

1) ZANA Technologies GmbH, Karlsruhe, Germany

2) Department of Respiratory Medicine, Institute of Nutrition and Translational Research in Metabolism (NUTRIM), Maastricht University, Netherlands

Presentner: prashanth@zana.ai

Motivation: Speech carries rich physiological signals that can be used to monitor a patient's health status non-invasively. In recent years, voice-based technologies have shown promising results in detecting symptoms related to chronic conditions such as Chronic Obstructive Pulmonary Disease (COPD) and Heart Failure (HF). However, most research in this area focuses narrowly on model performance, while often neglecting the broader infrastructure needed to deploy such systems in real-world settings. Zana offers an AI-driven, device-agnostic voice biomarker platform (VOICE-BIOME) that leverages the human voice as a novel biomarker for health monitoring. The cloud-based VOICE-BIOME solution delivers objective and unbiased data on a patient's disease progression and treatment response. A key innovation is the longitudinal analysis of voice data to predict critical events, such as heart decompensation in Heart Failure patients or exacerbations in COPD/asthma patients. Our results demonstrate the effectiveness of our methods: In heart Failure detection, the VOICE-BIOME platform achieves over 81% sensitivity and 79% specificity in identifying decompensated HF using voice features alone. For COPD, 79.0% \pm 3.0% of exacerbations are detected at least three days before onset, and 85.1% \pm 1.7% are identified by the onset day or earlier using voice measures alone.

This session introduces VOICE-BIOME, a modular and adaptable voice biomarker platform for speech-based health monitoring that supports data collection, acoustic feature processing, continuous retraining, and results visualization—all integrated into a streamlined pipeline. Instead of centering on specific machine learning algorithms, we aim to showcase how the underlying The platform enables reproducibility, cross-study comparability, and seamless integration with clinical workflows. We hope to open a discussion on how the VOICE-BIOME platform can be customized for different research contexts and clinical needs. An Overview of the modular speech processing and monitoring pipeline is given in Figure 1.

Data and Tools: The session will demonstrate our VOICE-BIOME platform using an in-house dataset comprising voice recordings from patients with chronic conditions such as COPD and Heart Failure. The recordings include structured speech tasks like vowel prolongation and text reading, collected via a mobile application. To facilitate daily voice data collection, we use a Digital Companion that encourages regular voice input through structured tasks, seamlessly integrating into users' routines to promote adherence and ensure reliable data capture. The dataset consists of 204 participants, contributing a total of 298 hours of recordings across 41,034 speech samples. The participants are categorized into three cohorts: Chronic Obstructive Pulmonary Disease (COPD), Asthma, and Acute Heart Failure (AHF). The dataset includes 19,868 vowel recordings, 13,860 text readings, and 7,306 spoken answers, providing a diverse set of voice

data for analysis. Participants in the COPD and Asthma cohorts were tracked with daily voice recordings over a 3-month period, while those in the Heart Failure group were monitored during hospitalization with decompensated HF and for six months post-discharge.

Demographic distribution varies across cohorts, with a mix of male and female participants. In addition to speech recordings, the dataset includes questionnaire responses, such as EXACT, and physiological measurements, including weight, ejection fraction, and NT-proBNP levels. We will showcase how the VOICE-BIOME platform handles data ingestion, novel acoustic feature extraction, automated retraining, and visualization through a modular pipeline. All demonstrations will be conducted using preloaded data, and participants are not required to bring any datasets or tools of their own.

Learning Goals: By the end of the session, participants will gain:

- An understanding of speech-based feature extraction and machine learning for health monitoring – including how acoustic features relevant to respiratory and cardiovascular conditions can be derived from recorded speech and used for exacerbation detection.
- Familiarity with a reusable end-to-end framework for speech analysis – learning how a modular pipeline handles data ingestion, feature processing, automated model retraining, prediction storage, and clinical dashboard integration.

Preparations: Participants do not need to bring anything to the session. All necessary data, tools, and demonstrations will be provided and presented live. No prior setup, software installation, or data preparation is required.

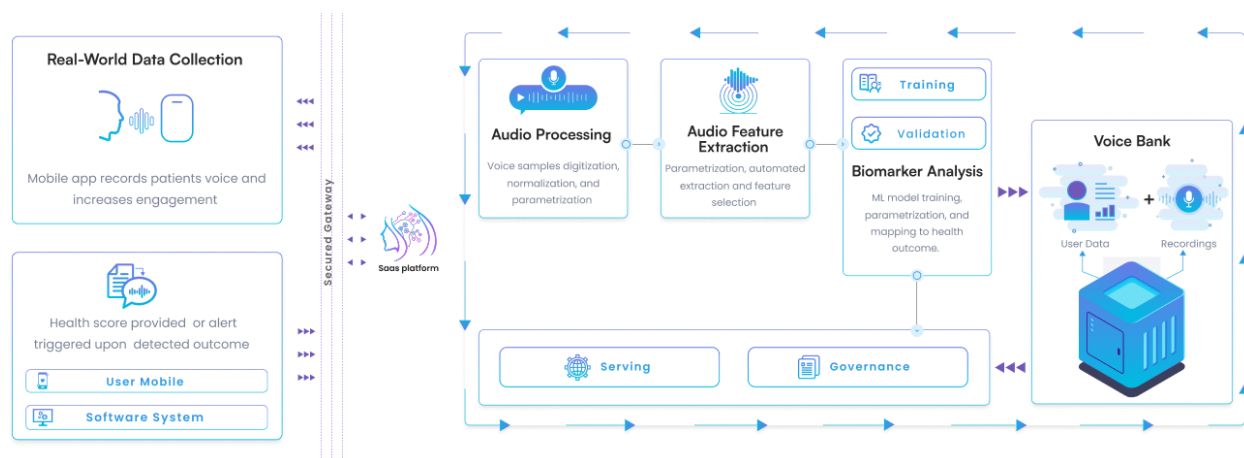


Figure 1: End-to-End Speech-Based Health Monitoring Framework: A scalable platform for collecting, processing, and analyzing patient speech data to detect health conditions. The system integrates real-world data collection, feature extraction, biomarker analysis, and secure data management to provide actionable health insight