Database design for RESONANZ

Original date: 2025-07-15 Updated: 2025-07-29

Author: Matthew Beaudoin

1 Introduction

The following is a specification for the database design for the RESONANZ project. Specifically, this specification is for the processed (segmented) audio data and assumes that the source data has been split from its original format to individual files separated by original speaker. The method used to perform this segmentation is not covered in this document, but will also be written down (at some point).

2 Specification

Audio data is recorded from various speakers in laboratory, clinical, or crowdsourced contexts. The audio data is separated into *segments*. A segment is a recording of speech consisting of only the voice of the speaker in question. Segments must have a duration in the range 4–12 seconds. The recording need not only contain natural breakages or pauses in speech; in other words, a speaker may be cut off or interrupted by the end of the file.

2.1 Data labels

The mandatory labels (column names) for the database, as well as there shortened key names, are as follows:

`db` a descriptive name of the original context or experiment from which the data was taken.

`lang` an ISO 639 two-letter language code (e.g., `de`, `en`) indicating the spoken language.

`pid` a unique (to the `db`/database) numeric identifier, indiciating a specific speaker.

`snum` a unique (to the `pid`/speaker) numeric identifier, indicating a specific audio segment.

`sex` a single-letter code indicating the sex of the speaker; possible options are `m` (male), `f` (female), `d` (Divers), and `u` (unknown/not shared).

`phq-9` nine-item Patient Health Questionnaire (PHQ-9) score assessing the severity of depressive symptoms; must take on an integer value between 0 and 27, inclusive (`phq-9` $\in [0, 27]$); zero indicates the lowest severity.

2.2 Optional data labels

`file_path` a relative path to the segment file according to the scheme `./{db}/{pid}/{file_num}.<extension>`.

`gad-7` Seven-item Generalized Anxiety Disorder (GAD-7) score assessing the severy of generalized anxiety disorder; must take on an integer value between 0 and 21, inclusive (`gad-7` $\in [0, 21]$); zero indicates the lowest severity.

3 Sample table

The database table will be stored as a `csv` file.

db	lang	pid	file_num	file_path	sex	phq-9	gad-7
crowdee	de	0001	0001	crowdee/0001/s0001.wav	f	8	4
crowdee	de	0001	0002	crowdee/0001/s0002.wav	f	8	4
crowdee	de	0002	0001	crowdee/0001/s0001.wav	d	3	-
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮

db	lang	pid	file_num	file_path	sex	phq-9	gad-7
crowdee	de	n	0001	crowdee/n/s0001.wav	m	12	6
crowdee	de	n	0002	crowdee/n/s0002.wav	m	12	6
qulab	de	0001	0001	qulab/0001/s0001.wav	d	21	-
qulab	de	0001	0002	qulab/0001/s0002.wav	d	21	-
qulab	de	0001	0003	qulab/0001/s0003.wav	d	21	-
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
qulab	de	n	0001	qulab/<n>/s0001.wav	u	12	14

The `file_path` label is actually entirely unecessary, as it can be constructed as a relative path according to the scheme `./{db}/{pid}/{file_num}.wav`; the file extension must however be assumed.

However, it will be easier for everyone if the full file path is explicitly included.

4 Filesystem hierarchy

- data
  - crowdee
    - 0001
      - s0001.wav
      - s0002.wav
    - 0002
      - s0001.wav
    - …
    - meta
      - responses.csv
      - SurveyJS.json
      - …
  - qulab
    - 0001
      - s0001.wav
    - 0002
      - s0001.wav
      - s0002.wav
      - s0003.wav
    - …
    - meta
      - …
  - …