Introduction to OpenWillis

An open source python library for digital phenotyping

Digital Measurement of Mental Health Workshop
September 10, 2025

Michelle Worthington, PhD
Director of Science, Brooklyn Health

BrooklynHealth

BklynHlth, Inc.
195 Montague St, 14
Brooklyn, NY 11201
getintouch@brooklyn.health

Agenda

- Problem of measurement in psychiatry and clinical science
- The promise of digital phenotyping
- Introduction to OpenWillis
    - Why this library is different
    - What can OpenWillis measure
- Published examples of research using OpenWillis
- Getting started and resources for users
- Practical tips for data processing and analysis

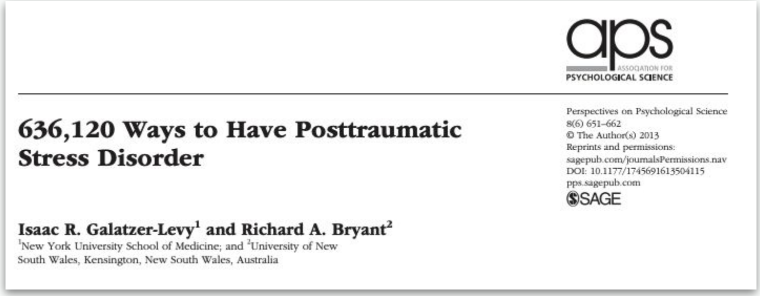
The measurement problem in psychiatry

(Brief review from earlier presentations)

Disorders manifest in observable behavior

Clinical assessments are meant to quantify behavioral manifestations of psychiatric conditions

This approach is limited due to significant heterogeneity in clinical presentations

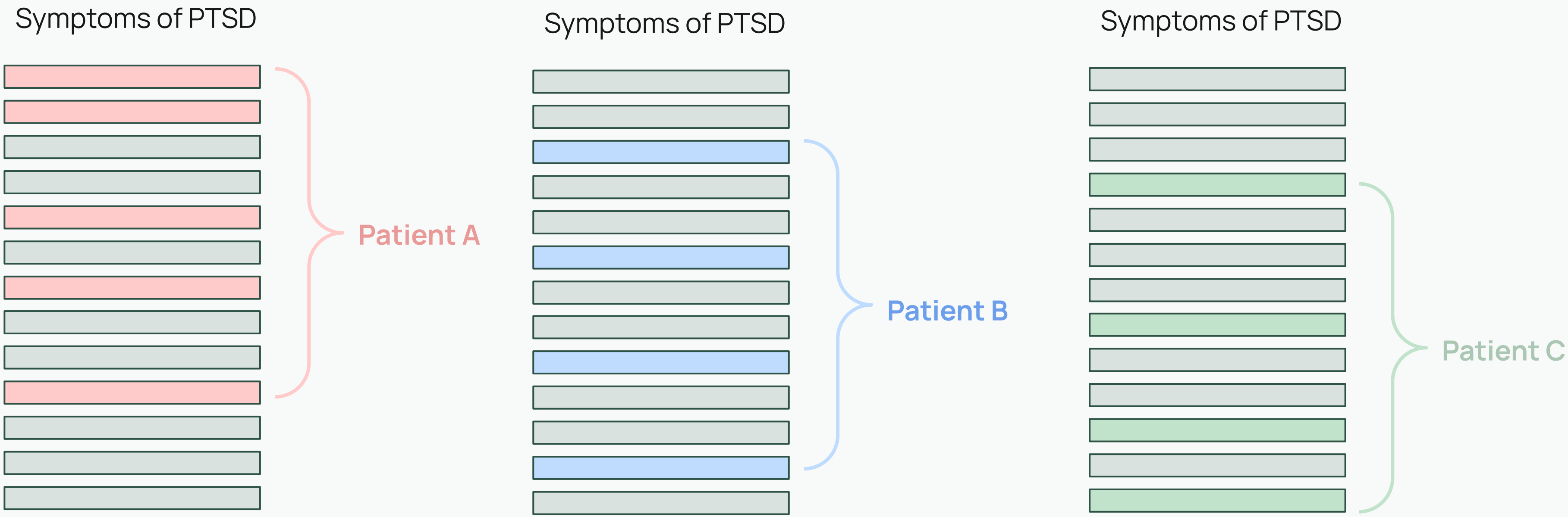


Perspectives on Psychological Science
8(6) 651–662

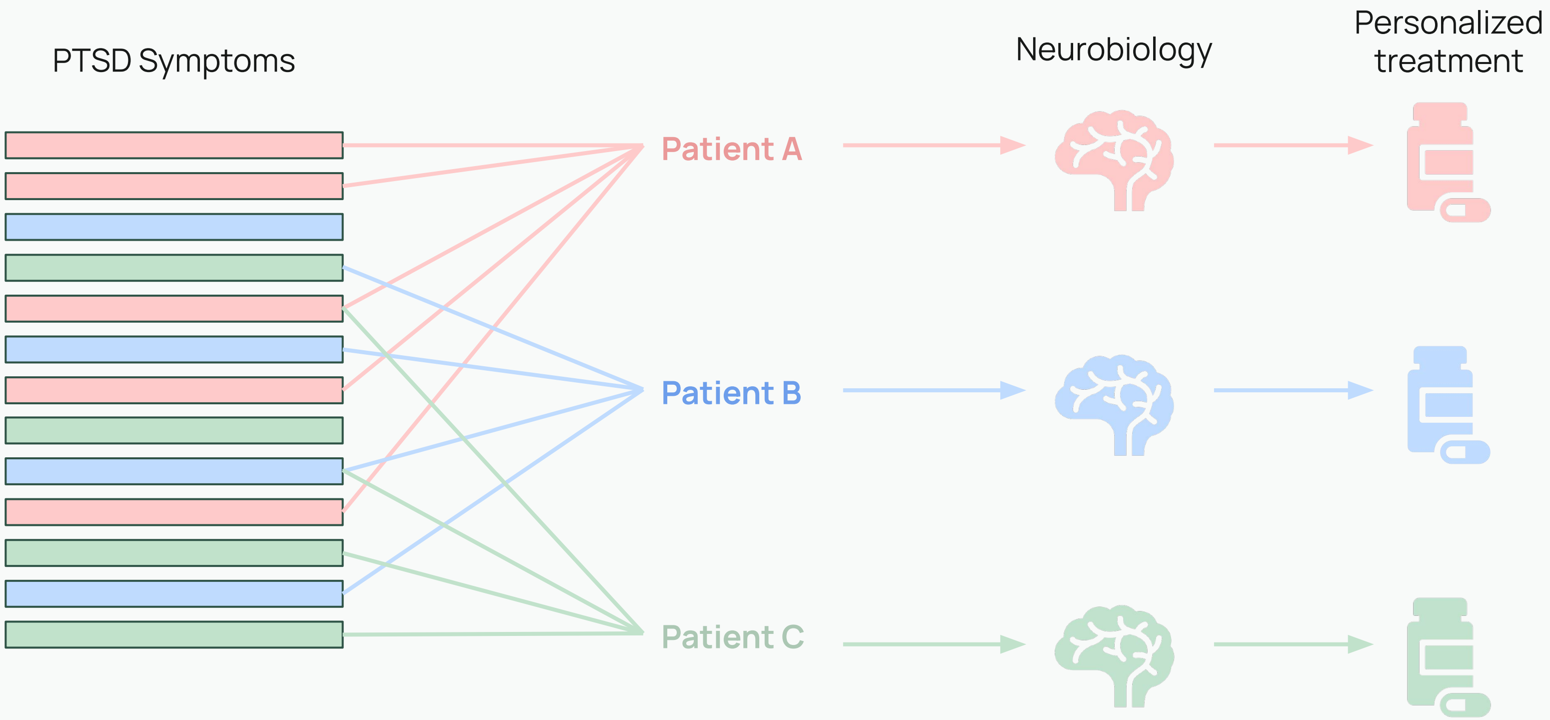
Reprints and permissions:
sagepub.com/journalsPermissions.nav
DOI: 10.1177/1745691613504115
pps.sagepub.com
SAGE

636,120 Ways to Have Posttraumatic Stress Disorder

Isaac R. Galatzer-Levy[1] and Richard A. Bryant[2]
[1]New York University School of Medicine; and [2]University of New South Wales, Kensington, New South Wales, Australia

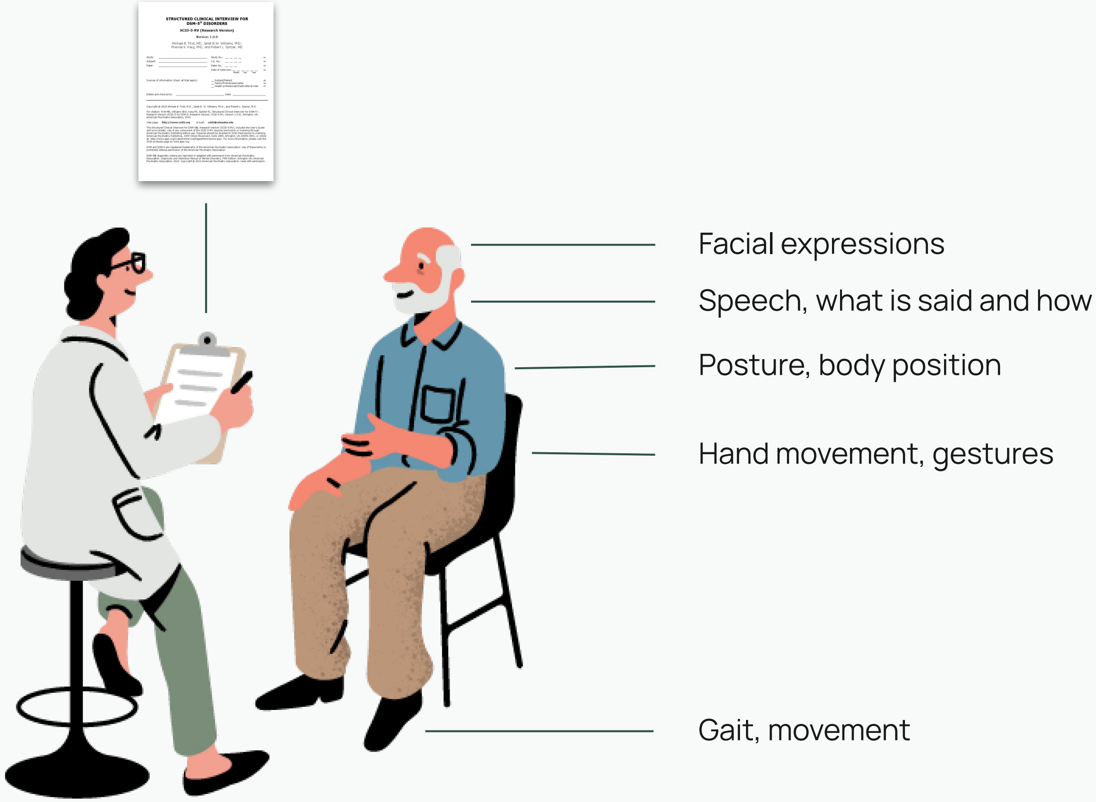


A possible future for mental health assessment and treatment:

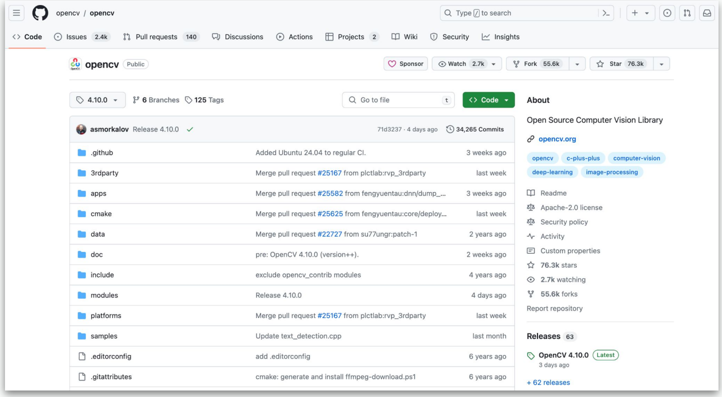


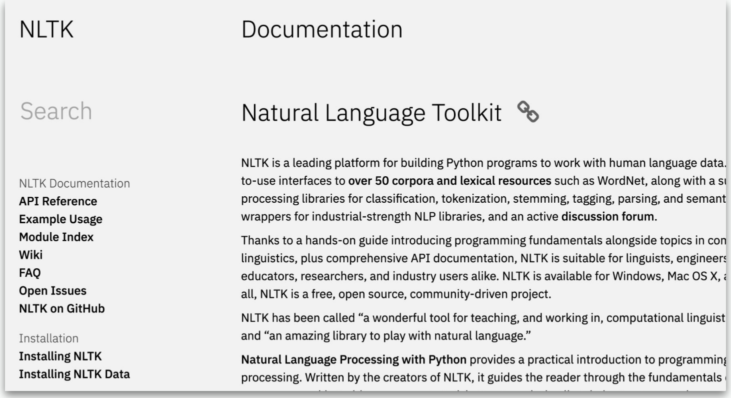
Digital health measures can enable precision measurement

Clinical interview administration



Digital phenotyping has been around for a while...

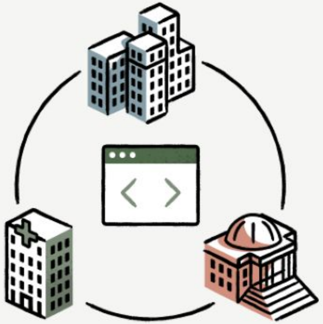




But the field has been lacking in a few areas:

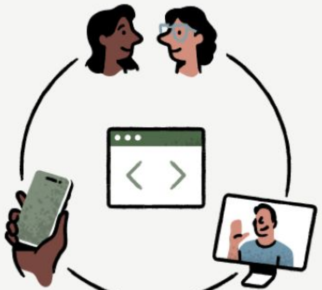
Trustworthy

Transparent methods and
open source code



Standardized

Same measures
regardless of context



Interoperable

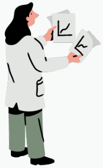
Platform-agnostic measures
serving all use cases

OpenWillis

An open source python library and repository for digital phenotyping

Why OpenWillis is different:

Centralized, open source repository for multimodal measures



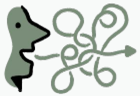
Ongoing methods validation in real world data

Users are encouraged to propose and contribute new methods

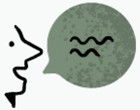
Active user community fosters collaboration and methods sharing

What OpenWillis can measure

Starting with a focus on audio and visual features, we are thoughtfully developing our bank of digital measures



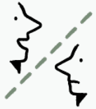
Vocal acoustics



Speech characteristics

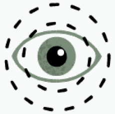
Speaker diarization

Speech transcription

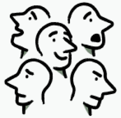


Facial expressivity

GPS analysis



Oculomotor analysis



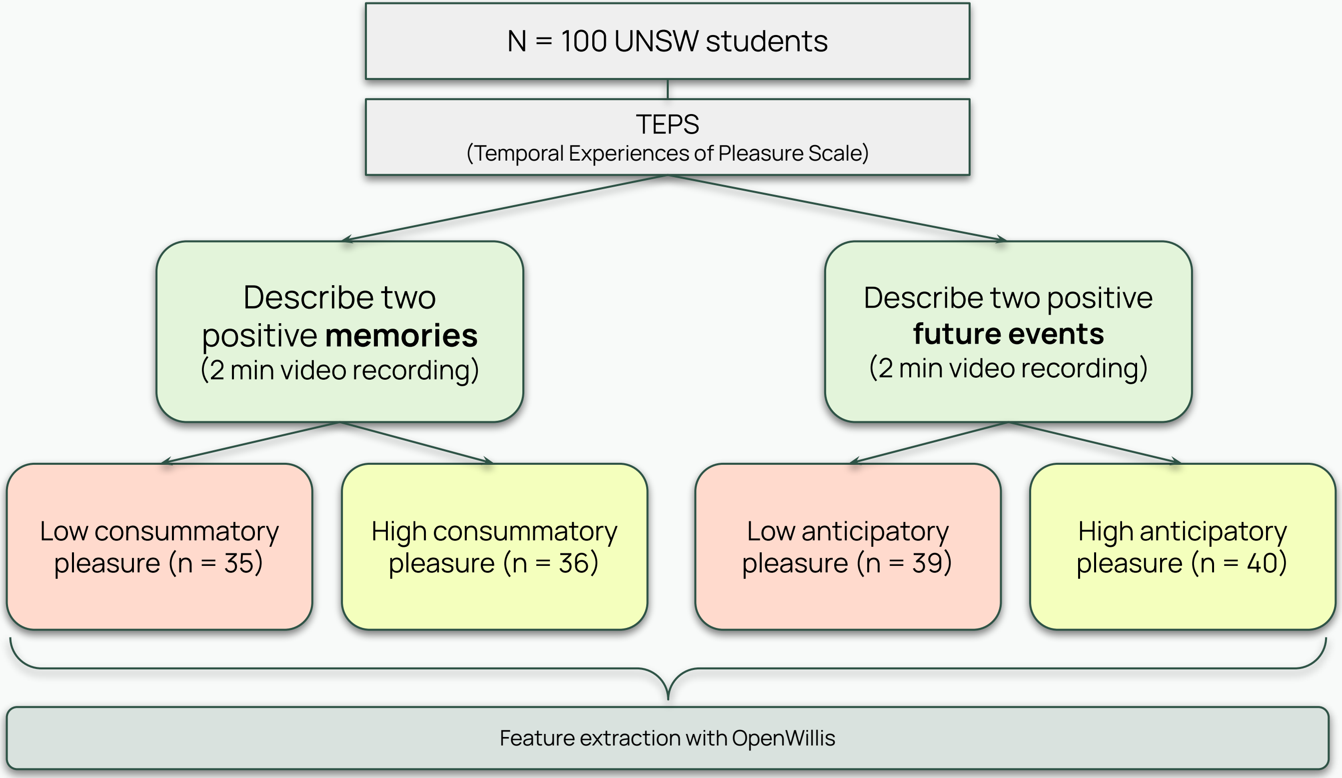
Emotional expressivity

Motor functioning

Case examples using OpenWillis methods

Example: Digital measurement of biomarkers of anhedonia

Aim: Identify digital biomarkers using OpenWillis related to consummatory and anticipatory pleasure



(Meaney et al. 2025, *Emotion*)

Example: Digital measurement of biomarkers of anhedonia (Meaney et al. 2025, *Emotion*)

Results:

Low consummatory pleasure				Low anticipatory pleasure			
Digital health measure	***P* value**	**Cohen's d**	**Direction***	**Digital health measure**	***P* value**	**Cohen's d**	**Direction***
Self-report positivity	< 0.001	1.59	**Reduced**	Self-report positivity	< 0.001	1.14	**Reduced**
Happiness facial expression	0.63			Happiness facial expression	< 0.001	0.63	**Reduced**
Sadness facial expression	0.70			Sadness facial expression	0.01	0.44	**Increased**
Overall facial activity	0.19			Overall facial activity	0.51		
Mean vocal pitch	0.17			Mean vocal pitch	0.005	0.51	**Reduced**
Vocal pitch deviation	0.21			Vocal pitch deviation	0.04	0.36	**Reduced**
Mean vocal volume	0.30			Mean vocal volume	0.10		
Jitter	0.89			Jitter	0.01	0.45	**Increased**
Shimmer	0.28			Shimmer	0.02	0.42	**Increased**
Proportion of silent frames	0.13			Proportion of silent frames	0.002	0.56	**Increased**
Speech sentiment	0.05			Speech sentiment	0.97		
Lexical diversity	0.56			Lexical diversity	0.08		

*Compared to high consummatory/anticipatory pleasure group using a paired samples t-test

(Meaney et al. 2025, *Emotion*)

Example: Digital measurement of biomarkers of anhedonia (Meaney et al. 2025, *Emotion*)

Conclusions:

- No differences using digital health measures for **consummatory pleasure**
- Differences observed in hypothesized direction for **anticipatory pleasure only**
- Distinct patterns for consummatory vs. anticipatory pleasure using digital biomarkers
- Provides evidence that anticipatory and consummatory pleasure are separate processes
- Gives insight into potential biomarkers for assessing diminished pleasure capacity in non-clinical populations

(Meaney et al. 2025, *Emotion*)

Example: Measuring schizophrenia symptom severity through PANSS recordings

A 5-week clinical trial for acute schizophrenia in adults

- A 5-week clinical trial for acute schizophrenia in adults
- Conducted PANSS interviews at each visit to assess symptoms
- Evaluated speech characteristics based on PANSS interviews to extract:
    - Features important for *characterizing* symptoms
    - Features important for *predicting* PANSS scores

(Worthington et al. 2025, *Frontiers in Psychiatry*)

Measurement of positive / negative symptoms

Speech measures extracted from PANSS, compared against PANSS scores

		Correlation with **positive** symptoms				Correlation with **negative** symptoms			
Behavior	**Measure**	**Study 1**	**Study 2**	**Study 3**	**Combined**	**Study 1**	**Study 2**	**Study 3**	**Combined**
Amount of speech	Speech length in minutes, adjusted	+	+	+	+	−	−	−	−
	Speech length in words, adjusted	+	+	+	+	−	−	−	−
	Mean turn length, minutes	+	+	+	+	−	−	−	−
	Mean turn length, words	+	+	+	+	−	−	−	−
	Speaker percentage	+	+	+	+	−	−	−	−
Rate of speech	Words per minute	+	+	+	+	−	−	−	−
	Syllables per minute	+	+	+	+	−	−	−	−
Pause characteristics	Word pause length, mean	−	−	−	−	+	+	+	+
	Word pause length, variability	−		−	−	+	+	+	+
	Mean pre-turn pause	−				+	+	+	+
Emotional valence	Positive sentiment	−			−		−		
	Negative sentiment	+			+	+	+	+	+
Lexical richness	Moving average type token ratio	−	−	−	−	+	+	+	+
		n = 825	*n* = 1,523	*n* = 1,134	*n* = 3,482	*n* = 825	*n* = 1,523	*n* = 1,134	*n* = 3,482

(Worthington et al. 2025, *Frontiers in Psychiatry*)

Measurement of disorganized thought, uncontrolled hostility, and anxiety/depression

Speech measures extracted from PANSS, compared against Marder Factor scores

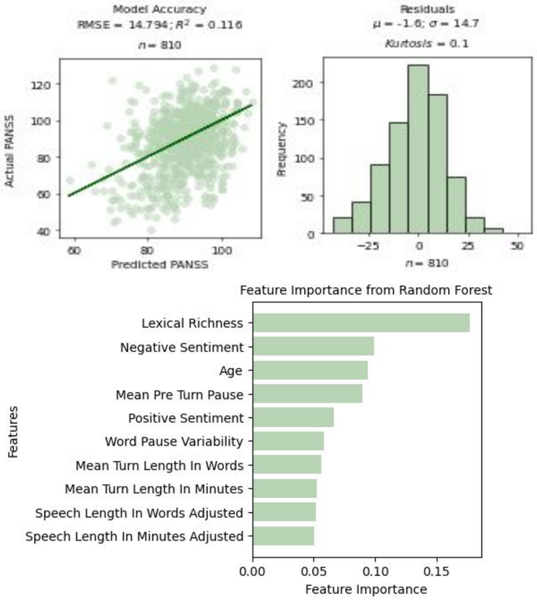
		Disorganized thought				Uncontrolled hostility				Anxiety / depression			
Behavior	**Measure**	**Study 1**	**Study 2**	**Study 3**	**Combined**	**Study 1**	**Study 2**	**Study 3**	**Combined**	**Study 1**	**Study 2**	**Study 3**	**Combined**
Amount of speech	Speech length in minutes, adjusted	+		+	+						+	+	+
	Speech length in words, adjusted	+	+	+	+	+		+	+	+	+	+	+
	Mean turn length, minutes	+	+	+	+			+	+		+	+	+
	Mean turn length, words	+	+	+	+	+		+	+	+	+	+	+
	Speaker percentage	+	+		+				+	+	+	+	+
Rate of speech	Words per minute					+	+	+	+	+		+	+
	Syllables per minute					+		+	+	+		+	+
Pause characteristics	Word pause length, mean							−		−		−	−
	Word pause length, variability									−			
	Mean pre-turn pause			+	+							+	+
Emotional valence	Positive sentiment				−	−	−		−		−	−	−
	Negative sentiment					+	+	+	+	+	+	+	+
Lexical richness	Moving average type token ratio	−		−	−	−	−	−	−	−	−	−	−
		n = 825	*n* = 1,523	*n* = 1,134	*n* = 3,482	*n* = 825	*n* = 1,523	*n* = 1,134	*n* = 3,482	*n* = 825	*n* = 1,523	*n* = 1,134	*n* = 3,482

(Worthington et al. 2025, *Frontiers in Psychiatry*)

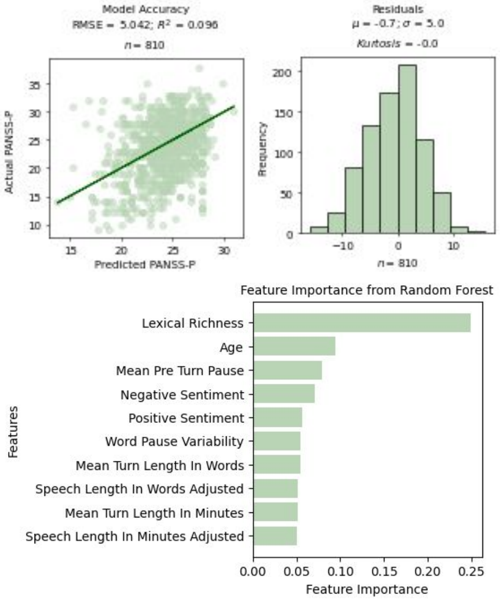
Model trained across studies

Prediction of PANSS scores from a model trained on speech measures

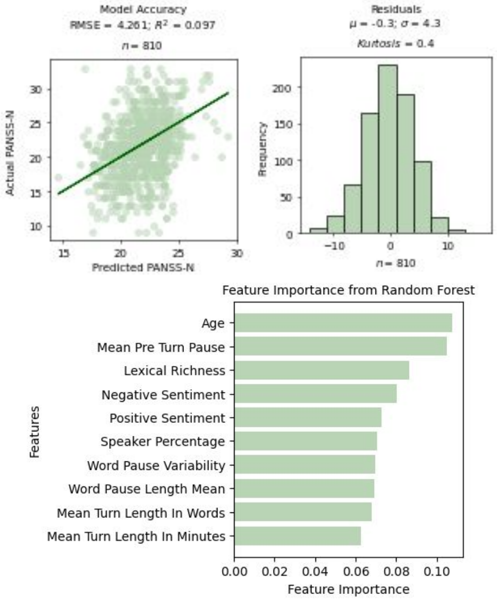
PANSS Overall



PANSS-P



PANSS-N



(Worthington et al. 2025, *Frontiers in Psychiatry*)

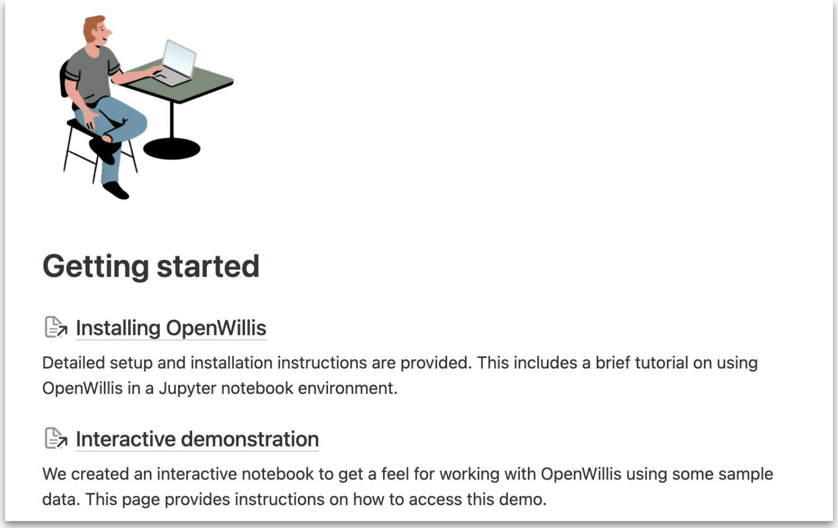
Getting started & user resources

Wiki: openwillis.brooklyn.health

GitHub: github.com/bklynhlth/openwillis

Installation

(For reference following this workshop!)



Getting started

Installing OpenWillis

Detailed setup and installation instructions are provided. This includes a brief tutorial on using OpenWillis in a Jupyter notebook environment.

Interactive demonstration

We created an interactive notebook to get a feel for working with OpenWillis using some sample data. This page provides instructions on how to access this demo.

OpenWillis Notion wiki home: openwillis.brooklyn.health

Tons of helpful resources can be found here

Option A – Complete installation

We highly recommend a dedicated virtual environment for OpenWillis. So, we'll start by creating this environment and installing the necessary dependencies.

Note: Throughout installation, if prompted to 'Proceed?' type 'y' and 'return' or 'enter'.

Step 1

Create and enter a virtual environment for OpenWillis. In a new terminal window, run the following:

```
conda create -n openwillis_env python=3.10

# Note: the full version of OpenWillis requires python version 3.9 or 3.10 even
though more recent versions of python are available

conda activate openwillis_env
```

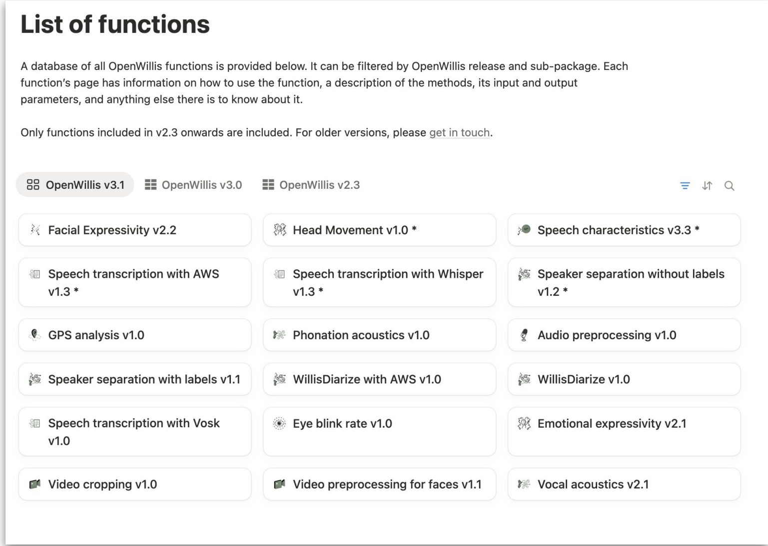
Step 2

After creating and activating the virtual environment in Step 1, proceed with installing the required packages `ffmpeg`, `portaudio`, `pysoundfile`, and `forest`:

Getting started › Installing OpenWillis

Contains instructions on how to install and use OpenWillis

List of functions and documentation



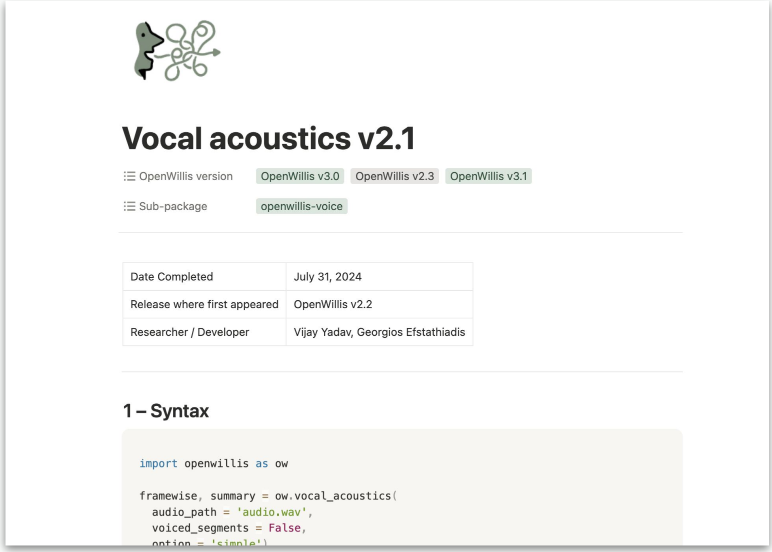
List of functions

A database of all OpenWillis functions is provided below. It can be filtered by OpenWillis release and sub-package. Each function's page has information on how to use the function, a description of the methods, its input and output parameters, and anything else there is to know about it.

Only functions included in v2.3 onwards are included. For older versions, please get in touch.

OpenWillis v3.1 | OpenWillis v3.0 | OpenWillis v2.3

Facial Expressivity v2.2	Head Movement v1.0 *	Speech characteristics v3.3 *
Speech transcription with AWS v1.3 *	Speech transcription with Whisper v1.3 *	Speaker separation without labels v1.2 *
GPS analysis v1.0	Phonation acoustics v1.0	Audio preprocessing v1.0
Speaker separation with labels v1.1	WillisDiarize with AWS v1.0	WillisDiarize v1.0
Speech transcription with Vosk v1.0	Eye blink rate v1.0	Emotional expressivity v2.1
Video cropping v1.0	Video preprocessing for faces v1.1	Vocal acoustics v2.1



Vocal acoustics v2.1

OpenWillis version: OpenWillis v3.0, OpenWillis v2.3, OpenWillis v3.1

Sub-package: openwillis-voice

Date Completed	July 31, 2024
Release where first appeared	OpenWillis v2.2
Researcher / Developer	Vijay Yadav, Georgios Efstathiadis

1 – Syntax

```
import openwillis as ow

framewise, summary = ow.vocal_acoustics(
  audio_path = 'audio.wav',
  voiced_segments = False,
  option = 'simple')
```

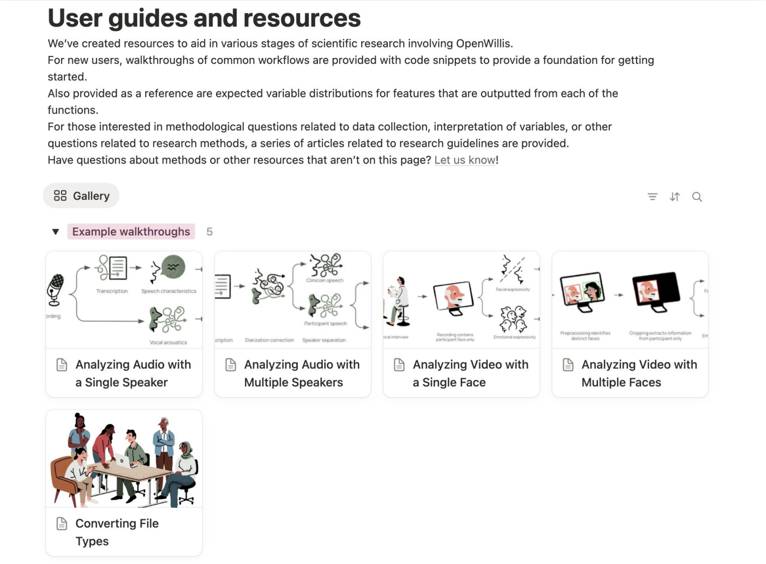
List of functions

Each function calculates a group of measures related to a modality

Function documentation

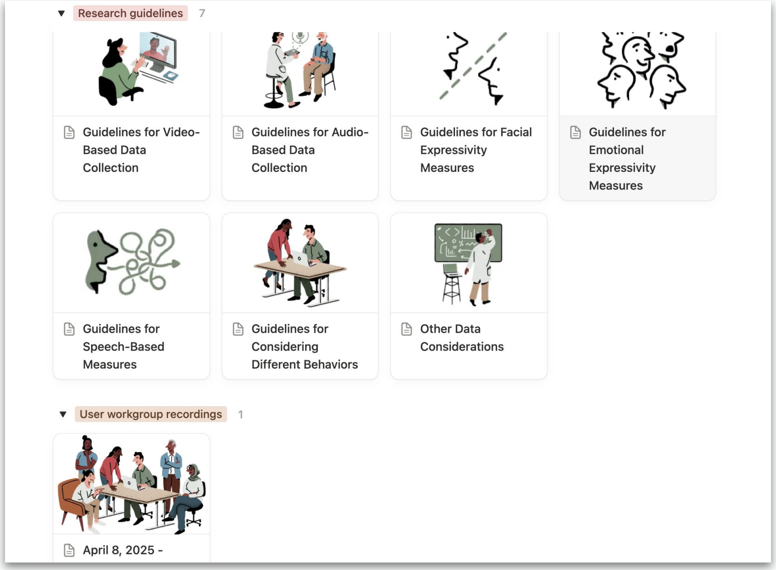
Instructions on use, methods utilized, and details on output variables found here

User resources and guidelines



Example walkthroughs

Provide additional guidance for data processing streams



Additional resources

Covers research guidelines, data considerations, etc.

Practical tips for data analysis

Practical tips for data analysis

Data preparation: Ensure your audio/video files are in the correct format (audio: .mp3 or .wav; video: .mp4 or .mov)

Converting File Types

Type of resource Example walkthroughs

Here, you will find some useful scripts for converting file types to make sure your data is compatible with OpenWillis. For audio, OpenWillis currently supports .wav and .mp3 files. For video, .mov and .mp4 files are supported.

For these conversions, we will use the `pydub` library which will is already part of the OpenWillis installation.

A few simple lines can convert file types within your scripts using this library:

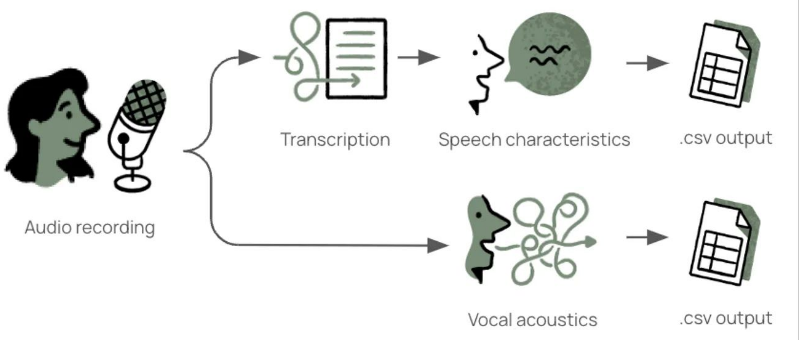
```
from pydub import AudioSegment

filepath = '/Users/researcher/project/data/audio_files/'

mp4_file = 'audio_sample.mp4' # specify existing file
output_file = 'audio_sample.wav' # creating a naming a new file

def convert_mp4_to_wav(input_file, output_file):
  try:
    audio = AudioSegment.from_file(input_file, format = "mp4")
    audio.export(output_file, format = "wav")
    print(f"Conversion successful! Saved as {output_file}")
  except Exception as e:
    print(f"An error occurred: {e}")
```

OpenWillis user resources contains a python script to convert your data into the correct format

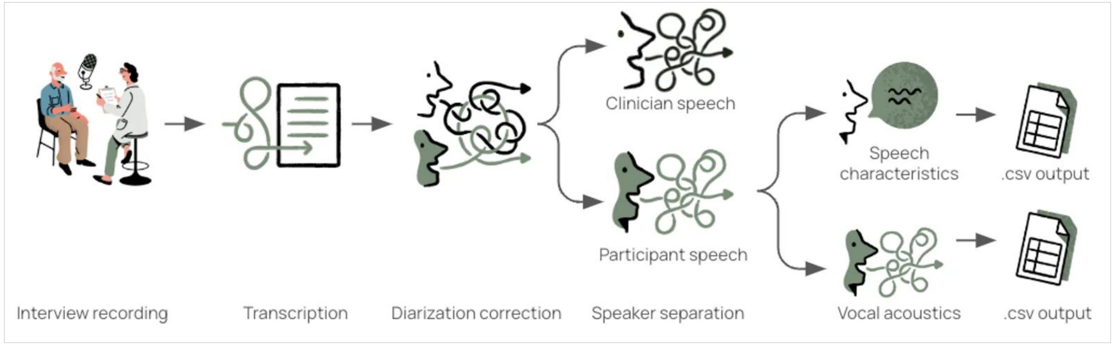
Practical tips for data analysis

Files with two speakers require extra steps for processing

Single speaker (audio):



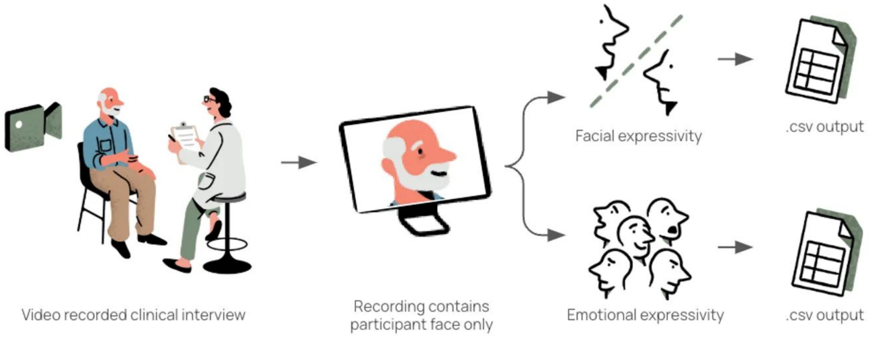
Multiple speakers (audio):



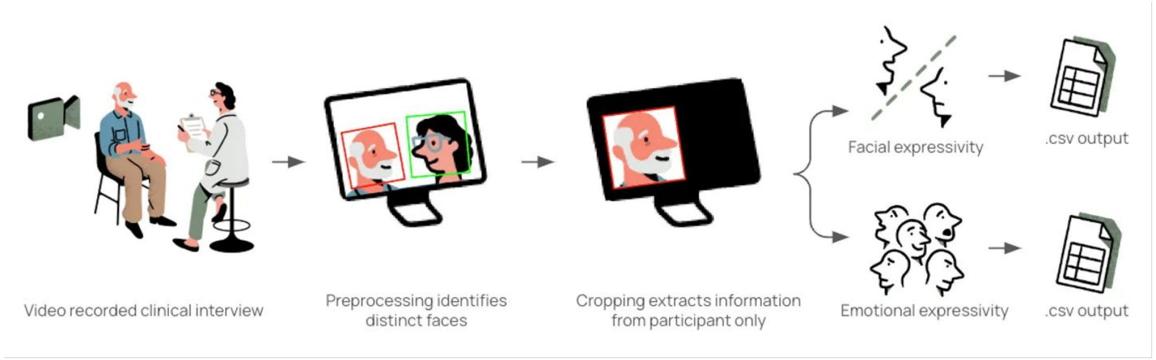
Practical tips for data analysis

Files with two speakers require extra steps for processing

Single speaker (video):



Multiple speakers (video):



Practical tips for data analysis

Function-specific considerations: **vocal acoustics**

1 – Syntax

```
import openwillis as ow

framewise, summary = ow.vocal_acoustics(
  audio_path = 'audio.wav',
  voiced_segments = False,
  option = 'simple')
```

3.3 – `option`

Type	String
Description	Default is 'simple', string that determines measures calculated; can be 'simple', 'advanced' or 'tremor'

Option	List of variables calculated
simple	Parselmouth measures, pause measures, cepstral measures
tremor	Simple measures + tremor measures
advanced	Simple measures + tremor measures and glottal measures

This function has the capability to output several dozen features related to voice, as we will see in the practical part of this workshop

Users have the option to output a **shorter** list of features when the parameter **`option`** is set to "**`simple`**"

Be sure to reference the documentation to determine the best choice for your research question

Practical tips for data analysis

Function-specific considerations: **speech characteristics**

2.1 – Per-word measures

The function's first output is a `words` dataframe, which contains a row for each word and measures specific to that word. This includes:

- `pre_word_pause`: pause time before the word in seconds. Individual word timestamps in the input JSON are used to calculate pause lengths before each word. To avoid measurement of potentially long silences prior to the start of speech in an audio file, the `pre_word_pause` for the *first* word in every file is set to `NaN`. To distinguish `pre_word_pause` from `pre_turn_pause` as defined later in this document, `pre_word_pause` for the first word in each turn is also set to `NaN`.
- `num_syllables`: number of syllables, identified using NLTK's SyllableTokenizer. This is an English-only measure.
- `part_of_speech`: part of speech associated with the word, identified using NLTK, as specified in the `part_of_speech` column (Andreasen & Pfohl, 1976; Tang et al., 2021); these are English only measures:
  - `noun`
  - `verb`
  - `adjective`
  - `pronoun`
  - `adverb`
  - `determiner`

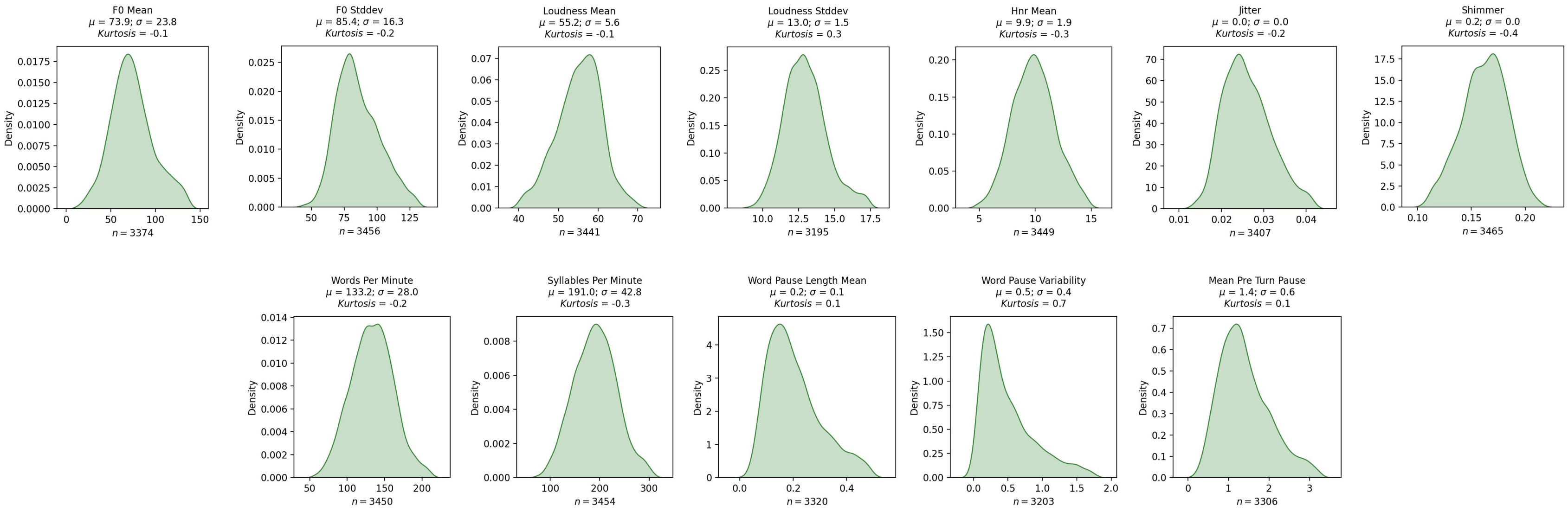
We are not looking at speech characteristics in this workshop because many of the features are not compatible with non-English languages

Features like speech rate, vocabulary tagging, and more complex NLP features are currently only available for English :(

We hope this changes in the future… in the meantime, always check the documentation to make sure certain features work for your language of interest

Practical tips for data analysis

Once your video/audio files are processed, be sure to examine the distributions of each variable



Some features tend to be skewed, such as pause length and variability

For even more recommendations, thoughts on data collection, and research guidelines...

Be sure to visit our wiki!

openwillis.brooklyn.health

For library updates, user community events, and more...

Sign up for the OpenWillis listserv!

Scan this QR code to sign up →

Or go to tinyurl.com/openwillis

To get in touch:

openwillis@brooklyn.health

michelle.worthington@brooklyn.health

Thank you!

michelle.worthington@brooklyn.health

BrooklynHealth

BklynHlth, Inc.
195 Montague St, 14
Brooklyn, NY 11201
getintouch@brooklyn.health