

Category	Full Name	Related Disease(s)	Language	Data Scale	Identity	Size	Institution	Data Availability	URL
Voice Disorder	VOICED Database	Various voice disorders	Italian	208 participants (150 pathological individuals and 58 healthy individuals)	Gender, Age, Lifestyle habits, Occupation	110.1MB	School of Medicine, University of Naples "Federico II"	Publicly available	https://physionet.org/content/voiced/1.0.0/
	Saarbruecken Voice Database (SVD)	Various pathologies	German	2225 participants (869 healthy individuals, 1356 pathological individuals)	Gender, Age	Total 38.1GB	Saarland University, Germany	Publicly available	https://stimmdb.coli.uni-saarland.de/
Neurodegenerative	CMU PITT Corpus (CMU PITT)	Alzheimer’s Disease, Dementia	English	397 participants (104 healthy individuals, 208 individuals, 85 unknown individuals)	Gender, Age	Total 17.7GB	Carnegie Mellon University, University of Pittsburgh	Publicly available	https://talkbank.org/dementia/access/English/Pitt.html
	Parkinson's Speech with Multiple Types of Sound Recordings (PD Database)	Parkinson's Disease	Turkish	252 participants (188 PWP, 64 healthy individuals)	Gender, Age	2MB	Department of Neurology in CerrahpaÅŸa Faculty of Medicine, Istanbul University	Publicly available	https://archive.ics.uci.edu/dataset/470/parkinson+s+disease+classification
	Alzheimer's Dementia Recognition through Spontaneous Speech (ADReSSo)	Alzheimer's Disease	English	AD Classification Task: 237 participants (166 training + 71 test) ; Cognitive Decline Task: 105 recordings	Gender, Age	/	ADReSS Challenge Organizers	First join as a DementiaBank member (free).	https://talkbank.org/dementia/ADReSSo-2021/index.html
Mental & Emotional Health	Distress Analysis Interview Corpus (DAIC-WOZ)	Depression, Anxiety	English	189 participants (47 depressed individuals, 142 healthy individuals)	Gender, Age	Total 52.82GB	University of Southern California (USC) Institute for Creative Technologies	Publicly available	https://dcapswoz.ict.usc.edu/
	ASDBank Dutch Asymmetries Corpus	Autism Spectrum Disorder (ASD)	Dutch	192 participants (109 healthy individuals, 83 pathological individuals)	Gender, Age	152MB	University of Groningen	Publicly available	https://talkbank.org/asd/access/Dutch/Asymmetries.html
Speech, Language & Developmental	TORGO Database	Dysarthria, Cerebral Palsy (CP), Amyotrophic Lateral Sclerosis (ALS)	English	15 speakers (8 with dysarthria, 7 healthy controls)	Speaker ID, Gender, Age	18GB (uncompressed)	University of Toronto, Holland Bloorview Kids Rehab Hospital	Publicly available	https://www.cs.toronto.edu/~complingweb/data/TORGO/torgo.html
	UASpeech	cerebral palsy	English	19 pathological individuals	Gender, Age	/	University of Illinois	Publicly available	https://speechtechnology.web.illinois.edu/uaspeech/
Infectious & Respiratory	Cambridge COVID-19 Sound Database (CSSD)	COVID-19, Asthma, Cough Multilingual (Crowdsourced)	Multilingual	53,449 audio samples (over 552 hours in total) crowd-sourced from 36,116 participants (2,106 samples	Gender, Age, Smoker status	/	University of Cambridge	Publicly available, DTA needs to be signed	https://www.covid-19-sounds.org/zh/blog/neurips_dataset.html

Pathological Voice Datasets

Main uses	Full Name	Language	Data Scale	Identity Attributes	Institution	Data Availability	URL
ASR (Automatic Speech Recognition)	LibriSpeech ASR Corpus (LibriSpeech)	English	~1000 hours	Speaker ID, Gender	Johns Hopkins University	Publicly available	https://www.openslr.org/12/
	Mozilla Common Voice (Common Voice)	Multilingual	>30,000 hours	Gender, Age group, Accent	Mozilla Foundation	Publicly available	https://commonvoice.mozilla.org/datasets
	Switchboard Corpus (Switchboard)	English	~260 hours, >500 speakers	Gender, Age group, Education level, Dialect region	Texas Instruments, LDC	Publicly available	https://catalog.ldc.upenn.edu/LDC97S62
	Open-Source Mandarin Speech Corpus (AISHELL-1)	Chinese	178 hours	Speaker ID, Gender	Beijing Shell Technology Co., Ltd.	Publicly available	https://www.openslr.org/33/
SER (Speech Emotion Recognition)	Interactive Emotional Dyadic Motion Capture Database (IEMOCAP)	English	12 hours of audiovisual data, including video, speech, motion capture of face, text transcriptions.	Speaker ID, Gender	University of Southern California (USC) - Signal Analysis and Interpretation Laboratory	Publicly available	https://sail.usc.edu/iemocap/
	Ryerson Audio-Visual Database of Emotional Speech and Song (RAVDESS)	English	7,356 files (24 professional actors)	Speaker ID, Gender, Emotion	Toronto Metropolitan University	Publicly available	https://zenodo.org/records/1188976
TTS (Text-to-Speech)	The LJ Speech Dataset (LJSpeech)	English	~24 hours (13,100 short audio clips)	Speaker ID, Gender	Curated by Keith Ito (from LibriVox public domain data)	Publicly available	https://keithito.com/LJ-Speech-Dataset/
	CSTR VCTK Corpus (Voice Cloning Toolkit)	English	~44 hours (109 native English speakers, ~400 sentences per speaker)	Speaker ID, Gender, Accent	University of Edinburgh	Publicly available	https://datashare.ed.ac.uk/handle/10283/2651
Speaker Recognition	VoxCeleb Large-Scale Speaker Recognition Datasets (VoxCeleb 1 & 2)	English	>2000 hours, >7000 speakers	Gender (Age and Nationality can be inferred from external sources)	University of Oxford	Publicly available	http://www.robots.ox.ac.uk/~vgg/data/voxceleb/
Acoustic-Phonetic	TIMIT Acoustic-Phonetic Continuous Speech Corpus (TIMIT)	English	~5 hours, 630 speakers	Gender, Age, Ethnicity, Dialect region, Education level	Texas Instruments, SRI, MIT	Publicly available	https://catalog.ldc.upenn.edu/LDC93s1

General-Purpose Voice Datasets