

VoicePrivacy 2026 Challenge

List of models, datasets and software allowed for training anonymization systems

version 12.05.2026

This table lists the models, datasets, and software permitted for the development and training of anonymization systems in the VoicePrivacy 2026 Challenge (VPC2026). Entries 1–15 (models), 76–91 (datasets) and 111–115 (software) are inherited from VPC2024; all remaining entries are in VPC2026 and highlighted in blue.

Where a link points to a webpage listing multiple versions of a model, all versions available on that page may be used unless otherwise stated. Participants may use any existing software; if it relies on pretrained models, those models must appear in this table.

The IEMOCAP dataset is excluded from the list of training data because it is used in the evaluation (see entry 104, EmoBox).

As the training resources submission deadline has already passed, no new resources will be allowed for fairness. However, teams are still welcome to explore unrequested models for research purposes. You may use them for experiments, submit the resulting anonymized data, and include the systems in your papers. These systems will be treated as “non-official” and will not be included in the official ranking.

Models

#	Model	Link
1	WavLM Base and Large	https://github.com/microsoft/unilm/tree/master/wavlm
2	Whisper	https://github.com/openai/whisper
3	HuBERT	https://github.com/facebookresearch/fairseq/blob/main/examples/hubert
4	XLS-R	https://github.com/facebookresearch/fairseq/blob/main/examples/wav2vec/xlsr
5	wav2vec 2.0	https://github.com/facebookresearch/fairseq/tree/main/examples/wav2vec https://dl.fbaipublicfiles.com/voxbopuli/models/wav2vec2_large_west_germanic_v2.pt
6	wav2vec2-large-robust-12-ft-emotion-msp-dim	https://huggingface.co/audenring/wav2vec2-large-robust-12-ft-emotion-msp-dim
7	ContentVec	https://github.com/auspicious3000/contentvec
8	w2v-BERT	https://github.com/facebookresearch/fairseq/tree/ust/examples/w2vbert
9	ECAPA2	https://huggingface.co/Jenthe/ECAPA2
10	ECAPA-TDNN	https://huggingface.co/speechbrain/spkrec-ecapa-voxceleb
11	NaturalSpeech 3	https://huggingface.co/amphion/naturalspeech3_facodec https://github.com/plachtaa/facodec
12	Hifi-GAN	Official implementation: https://github.com/jik876/hifi-gan NVIDIA Hifi-GAN Vocoder (en-US): https://huggingface.co/nvidia/tts_hifigan

#	Model	Link
13	CRDNN on CommonVoice 14.0 English	https://huggingface.co/speechbrain/asr-crdnn-commonvoice-14-en
14	Encodec	https://huggingface.co/facebook/encodec_24khz
15	Bark	https://huggingface.co/suno/bark https://huggingface.co/erogol/bark/tree/main
16	BigVGAN	https://github.com/NVIDIA/BigVGAN
17	Vocos	https://github.com/charactr-platform/vocos https://github.com/gemelo-ai/vocos
18	kNN-VC	https://github.com/bshall/knn-vc
19	FreeVC	https://github.com/OlaWod/FreeVC
20	SpeechT5	https://github.com/microsoft/SpeechT5
21	seed-vc	https://github.com/Plachtaa/seed-vc
22	OpenVoice	https://github.com/myshe11-ai/OpenVoice
23	ReDimNet	https://github.com/IDRnD/redimnet
24	TitaNet-Large	https://huggingface.co/nvidia/speakerverification_en_titanet_large
25	MFA-Conformer SV	https://github.com/zyzisyzy/mfa_conformer
26	ResNet-SE-34 (voxceleb_trainer)	https://github.com/clovaai/voxceleb_trainer
27	Canary-1B-Flash	https://huggingface.co/nvidia/canary-1b-flash
28	SeamlessM4T v2	https://github.com/facebookresearch/seamless_communication
29	emotion2vec	https://github.com/ddlBoJack/emotion2vec
30	IMS Toucan	https://github.com/DigitalPhonetics/IMS-Toucan
31	XTTS-v2	https://huggingface.co/coqui/XTTS-v2
32	Parler-TTS	https://github.com/huggingface/parler-tts
33	VibeVoice-1.5B	https://huggingface.co/microsoft/VibeVoice-1.5B
34	F5-TTS	https://github.com/SWivid/F5-TTS
35	AudioLDM 2	https://github.com/haoheliu/AudioLDM2
36	DiffWave	https://github.com/lmnt-com/diffwave
37	Grad-TTS / DiffVC	https://github.com/huawei-noah/Speech-Backbones/tree/main/Grad-TTS https://github.com/huawei-noah/Speech-Backbones/tree/main/DiffVC
38	NaturalSpeech 2	https://speechresearch.github.io/naturalspeech2
39	Voicebox-style flow matching	https://github.com/lucidrains/voicebox-pytorch
40	DAC (Descript Audio Codec)	https://github.com/descriptinc/descript-audio-codec https://huggingface.co/descript/dac_16khz

#	Model	Link
41	SNAC	https://github.com/hubertiuzdak/snac
42	Moshi	https://github.com/kyutai-labs/moshi
43	PESQ	https://github.com/ludlows/PESQ
44	STOI	https://github.com/mpariente/pystoi
45	ViSQOL	https://github.com/google/visqol
46	CREPE	https://github.com/maxrmorrison/torchcrepe
47	FCPE	https://github.com/CNChTu/FCPE
48	AutoVC	https://github.com/auspicious3000/autovc
49	Qwen2-Audio	https://huggingface.co/Qwen/Qwen2-Audio
50	ChatTTS	https://github.com/2noise/ChatTTS
51	StarGANv2-VC	https://github.com/yl4579/StarGANv2-VC
52	CosyVoice 2.0	https://huggingface.co/FunAudioLLM/CosyVoice2-0.5B
53	StyleTTS 2	https://github.com/yl4579/StyleTTS2
54	Matcha-TTS	https://github.com/shivammehta25/Matcha-TTS
55	VoxCPM2	https://github.com/OpenBMB/VoxCPM
56	UUVC	https://github.com/b04901014/UUVC
57	Vevo	https://github.com/open-mmlab/Amphion/blob/main/models/vc/vevo
58	Granite 4.0 1B Speech	https://huggingface.co/ibm-granite/granite-4.0-1b-speech
59	Qwen3-TTS	https://qwen.ai/blog?id=qwen3tts-0115 https://huggingface.co/collections/Qwen/qwen3-tts
60	CapSpeech	https://huggingface.co/OpenSound/CapSpeech-models
61	Fish-Audio	https://github.com/fishaudio/fish-speech https://huggingface.co/fishaudio/s2-pro
62	MeanVC	https://github.com/ASLP-lab/MeanVC
63	WeNet	https://github.com/wenet-e2e/wenet
64	Fast-U2++	https://github.com/cdliang11/fast-u2pp
65	SyllableLM	https://github.com/AlanBaade/SyllableLM
66	JHCodec	https://huggingface.co/jhcodec/jhcodec
67	SW2V	https://huggingface.co/jhcodec/sw2v_120k
68	StreamVoice (unofficial impl.)	https://github.com/hrnroh24/stream-vc
69	StreamVoiceAnon	https://github.com/Plachtaa/StreamVoiceAnon

#	Model	Link
70	WavLM-ECAPA (SSTC) ¹	https://duke.app.box.com/shared/static/na6grb7akap4ze66stiazp2azw4zb1f1
71	w2v-BERT 2.0	https://huggingface.co/facebook/w2v-bert-2.0
72	wav2vec2-large-960h	https://huggingface.co/facebook/wav2vec2-large-960h
73	Whisper-large-v3	https://huggingface.co/openai/whisper-large-v3
74	WavLM-ECAPA-joint	https://huggingface.co/espnet/voxcelebs12_ecapa_wavlm_joint
75	mHuBERT-147	https://huggingface.co/utter-project/mHuBERT-147

Datasets

#	Dataset	Link
76	ESD	https://hltsingapore.github.io/ESD/download.html
77	LibriSpeech: train-clean-100, train-clean-360, train-other-500	https://www.openslr.org/12
78	CREMA-D	https://github.com/CheyneyComputerScience/CREMA-D
79	RAVDESS	https://datasets.activeloop.ai/docs/ml/datasets/ravdess-dataset/ https://zenodo.org/records/1188976
80	VCTK	https://datashare.ed.ac.uk/handle/10283/3443
81	SAVEE	http://kahlan.eps.surrey.ac.uk/savee/ https://www.kaggle.com/datasets/ejlok1/surrey-audiovisual-expressed-emotion-savee
82	EMO-DB	http://emodb.bilderbar.info/download/
83	LJSpeech	https://keithito.com/LJ-Speech-Dataset/
84	Libri-light (only train part)	https://github.com/facebookresearch/libri-light/blob/main/data_preparation/README.md
85	VoxCeleb-1,2	https://www.robots.ox.ac.uk/~vgg/data/voxceleb/index.html#about
86	LibriTTS: train-clean-100, train-clean-360, train-other-500	https://openslr.org/60/
87	CMU-MOSEI	http://multicomp.cs.cmu.edu/resources/cmu-mosei-dataset/
88	MUSAN	https://www.openslr.org/17/
89	RIR	https://www.openslr.org/28/
90	VGAF (from EmotiW challenge)	https://sites.google.com/view/emotiw2023 https://www.kaggle.com/datasets/amirabdrhimov/vgaf-dataset
91	MSP-Podcast	https://ecs.utdallas.edu/research/researchlabs/msp-lab/MSP-Podcast.html

¹Same model as the lazy-informed ASV attacker; permitted for anonymization development but risks overfitting to this attacker.

#	Dataset	Link
92	LibriTTS-R	https://www.openslr.org/141/
93	LibriHeavy	https://huggingface.co/datasets/pkufool/libriheavy https://github.com/k2-fsa/libriheavy
94	MLS: Multilingual LibriSpeech (only train part)	https://www.openslr.org/94
95	VoxPopuli	https://github.com/facebookresearch/voxpathuli
96	GigaSpeech	https://github.com/SpeechColab/GigaSpeech
97	Hi-Fi TTS	https://www.openslr.org/109/
98	café	https://zenodo.org/record/1478765
99	PAVOQUE	https://github.com/marytts/pavoque-data
100	EmoV-DB	https://github.com/numediart/EmoV-DB
101	MEAD	https://github.com/uniBruce/Mead
102	TESS	https://doi.org/10.5683/SP2/E8H2ME
103	JL-Corpus	https://github.com/tli725/JL-Corpus
104	EmoBox (exclude IEMOCAP)	https://github.com/emo-box/EmoBox
105	Emilia	https://huggingface.co/datasets/amphion/Emilia-Dataset
106	WenetSpeech	https://github.com/wenet-e2e/WenetSpeech
107	Common Voice (English v23.0)	https://datacollective.mozillafoundation.org/organization/cmhf0j9o10006ns07jq45h7xk https://commonvoice.mozilla.org/
108	AISHELL-3	https://www.openslr.org/93
109	TidyVoice	https://mozilladatacollective.com/datasets/cmihstsewu023so207xot1iqqw
110	CAMEO	https://huggingface.co/datasets/amu-cai/CAMEO

Software with pretrained models

#	Software with pretrained models	Link
111	Resemblyzer	https://github.com/resemble-ai/Resemblyzer Model: https://github.com/resemble-ai/Resemblyzer/blob/master/resemblyzer/pretrained.pt
112	VITS	https://github.com/jaywalnut310/vits/ Models: https://drive.google.com/drive/folders/1ksarh-cJf3F5eKJlVWY0X1j1qsQqiS2
113	PIPER pretrained on VITS	https://github.com/rhasspy/piper/?tab=readme-ov-file Models: https://huggingface.co/datasets/rhasspy/piper-checkpoints/tree/main
114	RVC-Project	https://github.com/RVC-Project Models: https://huggingface.co/lj1995/VoiceConversionWebUI/tree/main

#	Software with pretrained models	Link
115	DISSC	https://github.com/gallilmaimon/DISSC
116	3D-Speaker	https://github.com/modelscope/3D-Speaker
117	HiFTNet	https://github.com/yl4579/HiFTNet
118	Masked prosody model	https://huggingface.co/cdminix/masked_prosody_model