VoicePrivacy 2024 Challenge: Final list of models and data for training anonymization systems version 26.03.2024

1. For models #1, 2, 3, 4, 5, 7, and 8, the provided link is a webpage listing multiple versions of the model. In this case, unless otherwise stated, all model versions available on that page before 21/03/2024 can be used by participants in the development and training of their anonymization systems.

2. Participants are allowed to use any existing software in the development and training of their anonymization systems. If the software uses pretrained models, these models should be explicitly listed in this table. This includes models #32-35 and the models listed on the main page (readme) of the repository #36 before 21/03/2024.

3. The IEMOCAP dataset is excluded from the list of training data because it is used in the evaluation.

4. The MSP-Podcast corpus is added to the list. For the purpose of the Challenge, the MSP-Podcast corpus providers can share the corpus for companies using the academic license. If a company wants to use the corpus beyond this challenge, it will have to obtain a commercial license by approaching the MSP-Podcast corpus providers. NEW

#	Model	Link
1	WavLM (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/voxpopuli/models/wav2vec2_large_west_germanic_v2.pt
6	wav2vec2-large-robust-12-ft-emotion-msp-dim	https://huggingface.co/audeering/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	NaturalSpeech3	https://huggingface.co/amphion/naturalspeech3_facodec
12	NVIDIA Hifi-GAN Vocoder (en-US)	https://huggingface.co/nvidia/tts_hifigan
13	CRDNN with CTC/Attention trained on CommonVoice 14.0 English (No LM)	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
#	**Dataset**	**Link**
16	ESD	https://hltsingapore.github.io/ESD/download.html
17	LibriSpeech: train-clean-100, train-clean-360, train-other-500	https://www.openslr.org/12
18	CREMA-D	https://github.com/CheyneyComputerScience/CREMA-D
19	RAVDESS	https://datasets.activeloop.ai/docs/ml/datasets/ravdess-dataset/ https://zenodo.org/records/1188976
20	VCTK	https://datashare.ed.ac.uk/handle/10283/2651 https://huggingface.co/datasets/vctk
21	SAVEE	http://kahlan.eps.surrey.ac.uk/savee/ https://www.kaggle.com/datasets/ejlok1/surrey-audiovisual-expressed-emotion-savee
22	EMO-DB	http://emodb.bilderbar.info/download/
23	LJSpeech	https://keithito.com/LJ-Speech-Dataset/
24	Libri-light (only train part)	https://github.com/facebookresearch/libri-light/blob/main/data_preparation/README.md
25	VoxCeleb-1,2	https://www.robots.ox.ac.uk/~vgg/data/voxceleb/index.html#about
26	LibriTTS: train-clean-100, train-clean-360, train-other-500	https://openslr.org/60/
27	CMU-MOSEI	http://multicomp.cs.cmu.edu/resources/cmu-mosei-dataset/
28	MUSAN noise	https://www.openslr.org/17/
29	RIR dataset	https://www.openslr.org/28/
30	VGAF dataset (from EmotiW challenge)	https://sites.google.com/view/emotiw2023 https://www.kaggle.com/datasets/amirabdrahimov/vgaf-dataset
31	MSP-Podcast NEW	https://ecs.utdallas.edu/research/researchlabs/msp-lab/MSP-Podcast.html
#	**Software with pretrained models**	**Link**
32	Resemblyzer	https://github.com/resemble-ai/Resemblyzer Model: https://github.com/resemble-ai/Resemblyzer/blob/master/resemblyzer/pretrained.pt
33	VITS	https://github.com/jaywalnut310/vits/ Models: https://drive.google.com/drive/folders/1ksarh-cJf3F5eKJjLVWY0X1j1qsQqiS2
34	PIPER pretrained on VITS	https://github.com/rhasspy/piper/?tab=readme-ov-file Models: https://huggingface.co/datasets/rhasspy/piper-checkpoints/tree/main
35	RVC-Project	https://github.com/RVC-Project Models: https://huggingface.co/lj1995/VoiceConversionWebUI/tree/main
36	DISSC	https://github.com/gallilmaimon/DISSC