

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

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 #31-34 and the models listed on the main page (readme) of the repository #35 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 might be added to the list within 10 days if it will be confirmed that it is free for companies for the purpose of the challenge.

#	Model	Link
1	WavLM (WavLM Base and Large)	<a href="https://github.com/microsoft/unilm/tree/master/wavlm">https://github.com/microsoft/unilm/tree/master/wavlm</a>
2	Whisper	<a href="https://github.com/openai/whisper">https://github.com/openai/whisper</a>
3	HuBERT	<a href="https://github.com/facebookresearch/fairseq/blob/main/examples/hubert">https://github.com/facebookresearch/fairseq/blob/main/examples/hubert</a>
4	XLS-R	<a href="https://github.com/facebookresearch/fairseq/blob/main/examples/wav2vec/xlsr">https://github.com/facebookresearch/fairseq/blob/main/examples/wav2vec/xlsr</a>
5	wav2vec 2.0	<a href="https://github.com/facebookresearch/fairseq/tree/main/examples/wav2vec">https://github.com/facebookresearch/fairseq/tree/main/examples/wav2vec</a> <a href="https://dl.fbaipublicfiles.com/voxpathli/models/wav2vec2_large_west_germanic_v2.pt">https://dl.fbaipublicfiles.com/voxpathli/models/wav2vec2_large_west_germanic_v2.pt</a>
6	wav2vec2-large-robust-12-ft-emotion-msp-dim	<a href="https://huggingface.co/audeering/wav2vec2-large-robust-12-ft-emotion-msp-dim">https://huggingface.co/audeering/wav2vec2-large-robust-12-ft-emotion-msp-dim</a>
7	ContentVec	<a href="https://github.com/auspicious3000/contentvec">https://github.com/auspicious3000/contentvec</a>
8	w2v-BERT	<a href="https://github.com/facebookresearch/fairseq/tree/ust/examples/w2vbert">https://github.com/facebookresearch/fairseq/tree/ust/examples/w2vbert</a>
9	ECAPA2	<a href="https://huggingface.co/Jenthe/ECAPA2">https://huggingface.co/Jenthe/ECAPA2</a>
10	ECAPA-TDNN	<a href="https://huggingface.co/speechbrain/spkrec-ecapa-voxceleb">https://huggingface.co/speechbrain/spkrec-ecapa-voxceleb</a>
11	NaturalSpeech3	<a href="https://huggingface.co/amphion/naturalspeech3_facodec">https://huggingface.co/amphion/naturalspeech3_facodec</a>
12	NVIDIA HiFi-GAN Vocoder (en-US)	<a href="https://huggingface.co/nvidia/tts_hifigan">https://huggingface.co/nvidia/tts_hifigan</a>
13	CRDNN with CTC/Attention trained on CommonVoice 14.0 English (No LM)	<a href="https://huggingface.co/speechbrain/asr-crdnn-commonvoice-14-en">https://huggingface.co/speechbrain/asr-crdnn-commonvoice-14-en</a>
14	Encodec	<a href="https://huggingface.co/facebook/encodec_24khz">https://huggingface.co/facebook/encodec_24khz</a>
15	Bark	<a href="https://huggingface.co/suno/bark">https://huggingface.co/suno/bark</a> <a href="https://huggingface.co/erogol/bark/tree/main">https://huggingface.co/erogol/bark/tree/main</a>
#	Dataset	Link
16	ESD	<a href="https://hltsingapore.github.io/ESD/download.html">https://hltsingapore.github.io/ESD/download.html</a>
17	LibriSpeech: train-clean-100, train-clean-360, train-other-500	<a href="https://www.openslr.org/12">https://www.openslr.org/12</a>
18	CREMA-D	<a href="https://github.com/CheyneyComputerScience/CREMA-D">https://github.com/CheyneyComputerScience/CREMA-D</a>
19	RAVDNESS	<a href="https://datasets.activeloop.ai/docs/ml/datasets/ravdess-dataset/">https://datasets.activeloop.ai/docs/ml/datasets/ravdess-dataset/</a> <a href="https://zenodo.org/records/1188976">https://zenodo.org/records/1188976</a>
20	VCTK	<a href="https://datashare.ed.ac.uk/handle/10283/2651">https://datashare.ed.ac.uk/handle/10283/2651</a> <a href="https://huggingface.co/datasets/vctk">https://huggingface.co/datasets/vctk</a>
21	SAVEE	<a href="http://kahlan.eps.surrey.ac.uk/savee/">http://kahlan.eps.surrey.ac.uk/savee/</a> <a href="https://www.kaggle.com/datasets/ejlok1/surrey-audiovisual-expressed-emotion-savee">https://www.kaggle.com/datasets/ejlok1/surrey-audiovisual-expressed-emotion-savee</a>
22	EMO-DB	<a href="http://emodb.bilderbar.info/download/">http://emodb.bilderbar.info/download/</a>
23	LJSpeech	<a href="https://keithito.com/LJ-Speech-Dataset/">https://keithito.com/LJ-Speech-Dataset/</a>
24	Libri-light (only train part)	<a href="https://github.com/facebookresearch/libri-light/blob/main/data_preparation/README.md">https://github.com/facebookresearch/libri-light/blob/main/data_preparation/README.md</a>
25	VoxCeleb-1,2	<a href="https://www.robots.ox.ac.uk/~vgg/data/voxceleb/index.html#about">https://www.robots.ox.ac.uk/~vgg/data/voxceleb/index.html#about</a>
26	LibriTTS: train-clean-100, train-clean-360, train-other-500	<a href="https://openslr.org/60/">https://openslr.org/60/</a>
27	CMU-MOSEI	<a href="http://multicomp.cs.cmu.edu/resources/cmu-mosei-dataset/">http://multicomp.cs.cmu.edu/resources/cmu-mosei-dataset/</a>
28	MUSAN noise	<a href="https://www.openslr.org/17/">https://www.openslr.org/17/</a>
29	RIR dataset	<a href="https://www.openslr.org/28/">https://www.openslr.org/28/</a>
30	VGAF dataset (from EmotiW challenge)	<a href="https://sites.google.com/view/emotiw2023">https://sites.google.com/view/emotiw2023</a> <a href="https://www.kaggle.com/datasets/amirabdrahimov/vgaf-dataset">https://www.kaggle.com/datasets/amirabdrahimov/vgaf-dataset</a>
#	Software with pretrained models	Link
31	Resemblyzer	<a href="https://github.com/resemble-ai/Resemblyzer">https://github.com/resemble-ai/Resemblyzer</a> Model: <a href="https://github.com/resemble-ai/Resemblyzer/blob/master/resemble-ai/pretrained.pt">https://github.com/resemble-ai/Resemblyzer/blob/master/resemble-ai/pretrained.pt</a>
32	VITS	<a href="https://github.com/jaywalnut310/vits/">https://github.com/jaywalnut310/vits/</a> Models: <a href="https://drive.google.com/drive/folders/1ksarh-cJf3F5eKJlLVWY0X1j1qsQqiS2">https://drive.google.com/drive/folders/1ksarh-cJf3F5eKJlLVWY0X1j1qsQqiS2</a>
33	PIPER pretrained on VITS	<a href="https://github.com/rhasspy/piper/?tab=readme-ov-file">https://github.com/rhasspy/piper/?tab=readme-ov-file</a> Models: <a href="https://huggingface.co/datasets/rhasspy/piper-checkpoints/tree/main">https://huggingface.co/datasets/rhasspy/piper-checkpoints/tree/main</a>
34	RVC-Project	<a href="https://github.com/RVC-Project">https://github.com/RVC-Project</a> Models: <a href="https://huggingface.co/lj1995/VoiceConversionWebUI/tree/main">https://huggingface.co/lj1995/VoiceConversionWebUI/tree/main</a>
35	DISSC	<a href="https://github.com/gallilmaimon/DISSC">https://github.com/gallilmaimon/DISSC</a>