

How do we make a speaking computer sound like anyone, at any age?

Easy, flexible, and controllable synthetic speech

どんな研究？

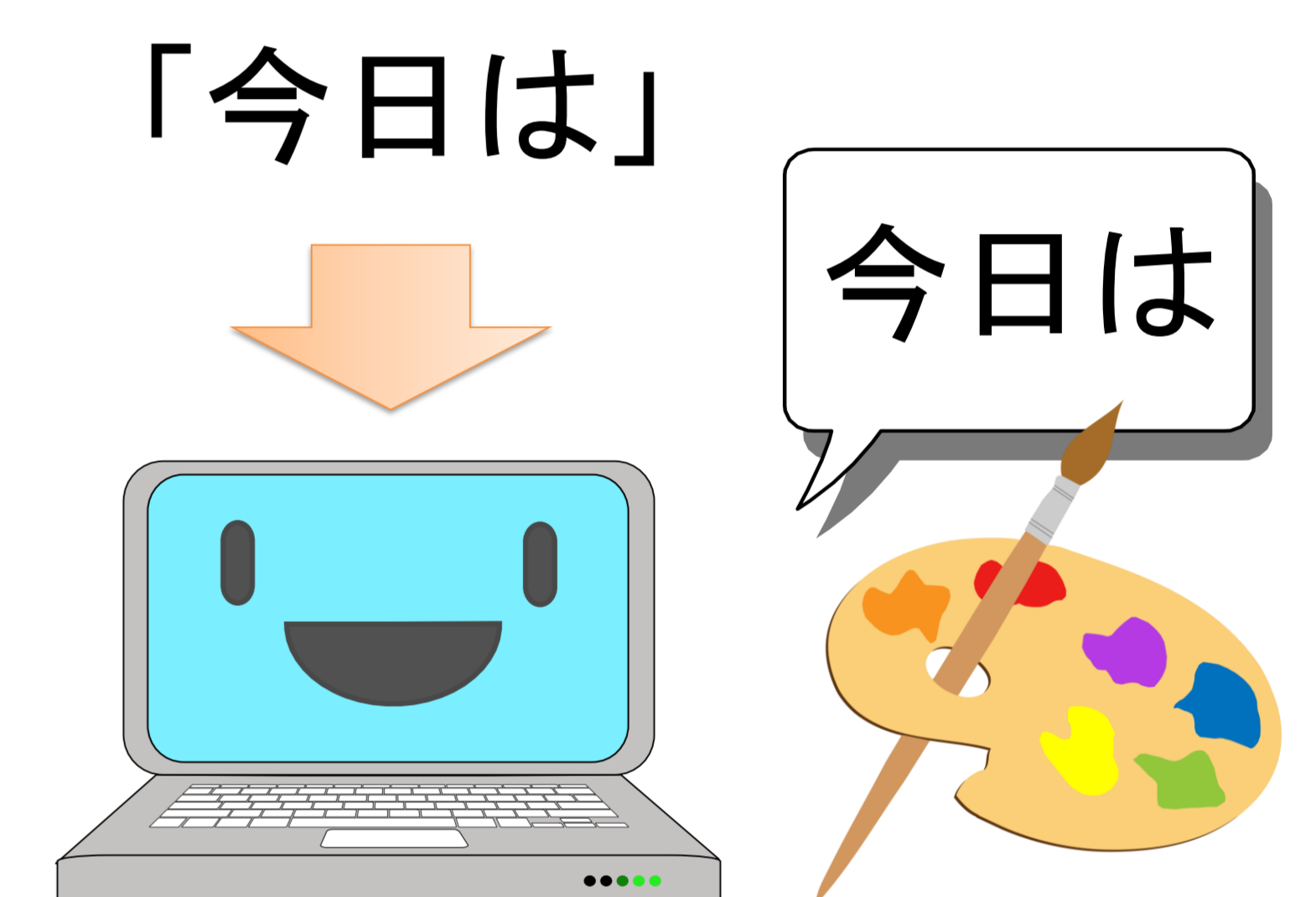
Gender, age, and anatomy make our voices sound different
Can computers learn this variation?
Can computers manipulate our voices as we wish?

何がわかる？

- DNNs can learn voice variation
- DNNs can mix several speakers and create new voices
- DNNs can age your voice
- Voice design

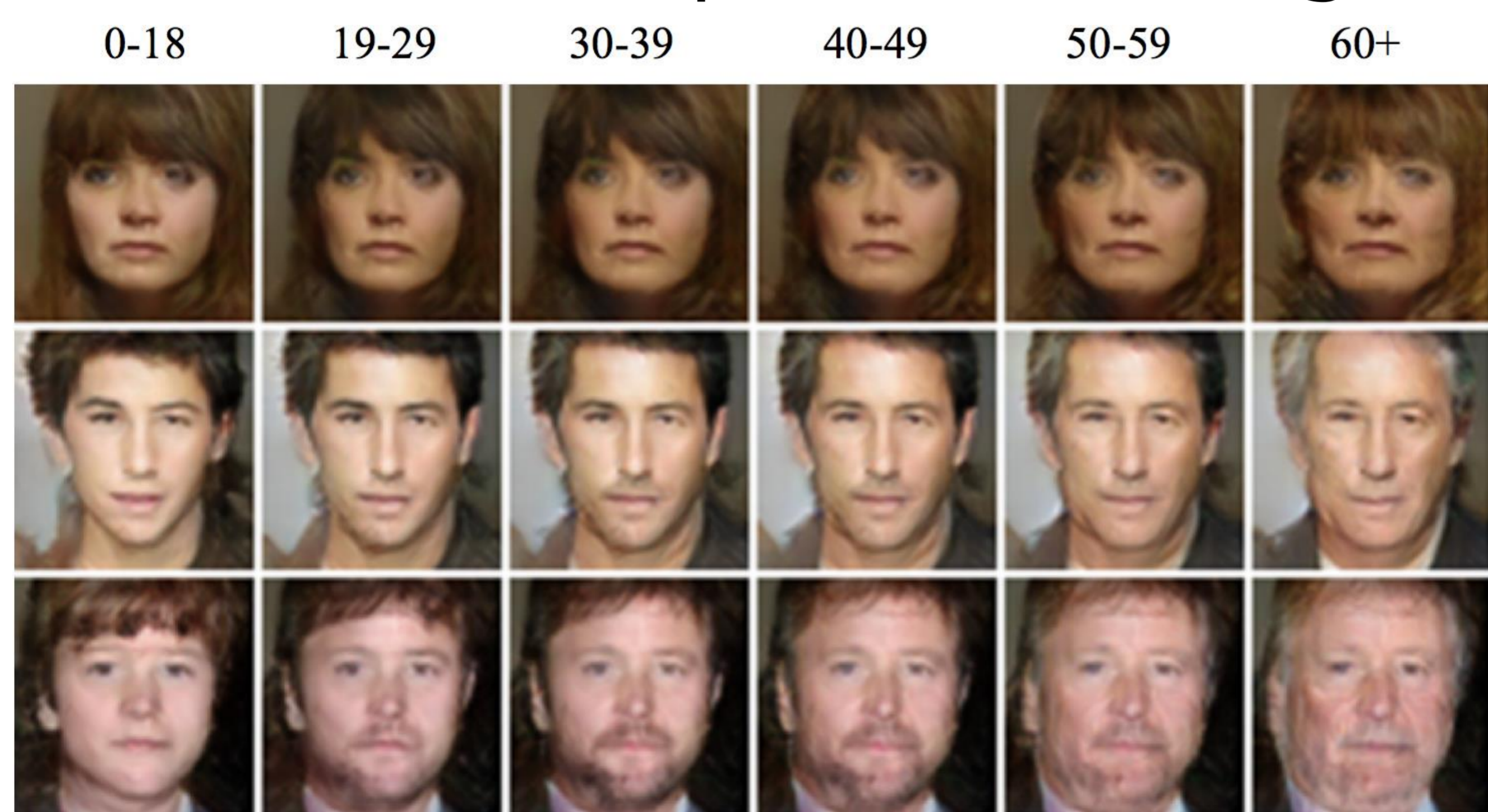
状況設定

- Record many speakers with different ages
- Train a DNN
- Generate speech with a specific age, gender, and speaker identity

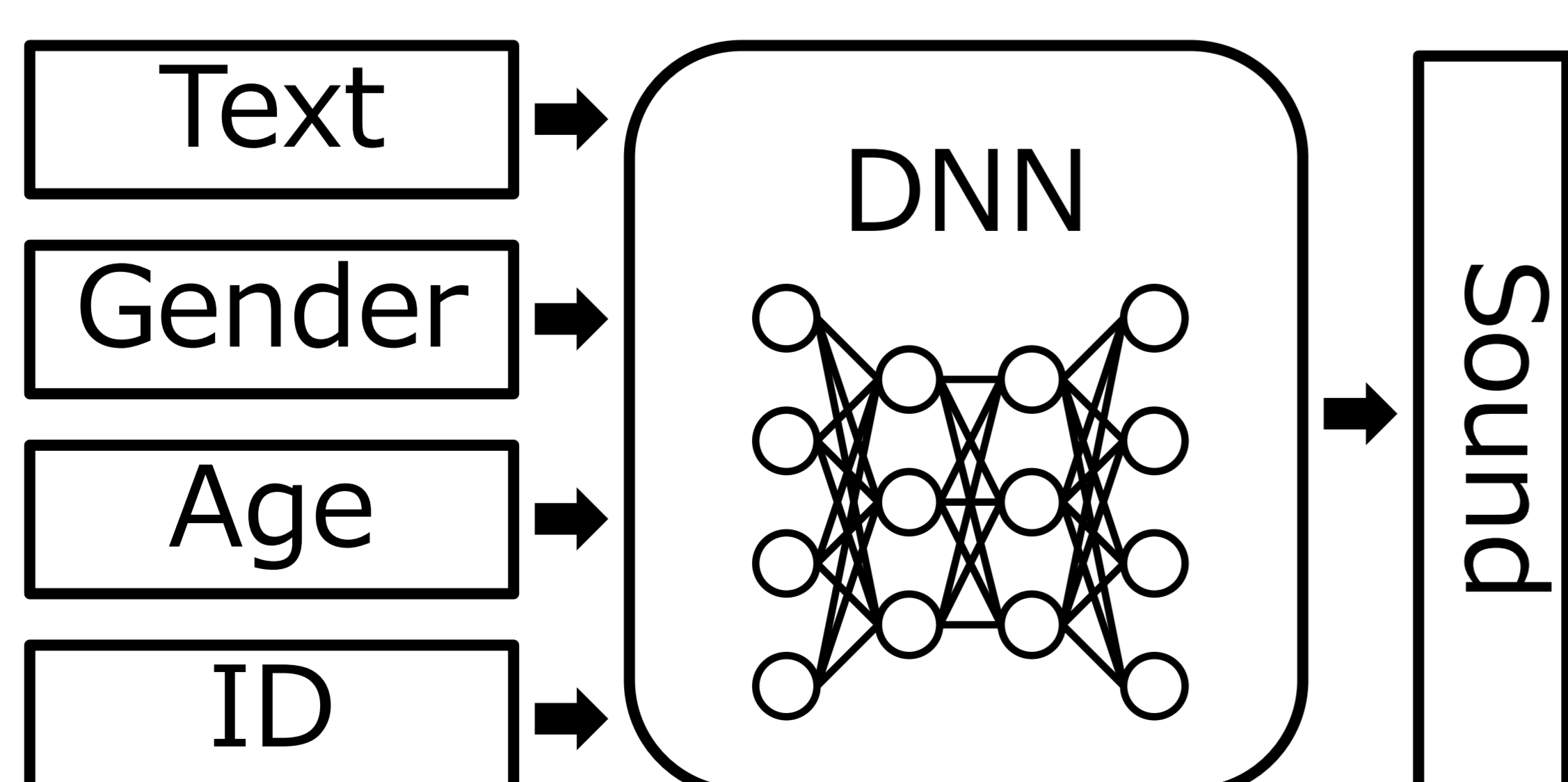


研究内容

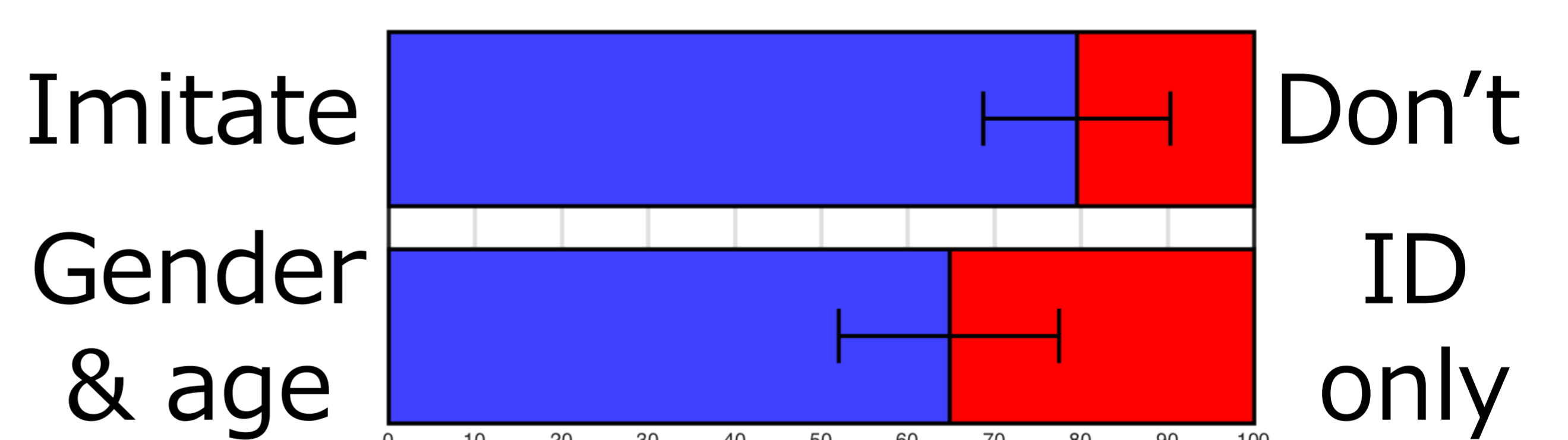
DNNs can manipulate facial age



- Requires many photos of people of different ages
- We want to change voice age!



56 males,
56 females,
100 utterances each
10~89 years



1. Can recreate training voices
2. Can imitate new voices
3. *Manipulation*: We can change speaker ID, gender, and age!

