

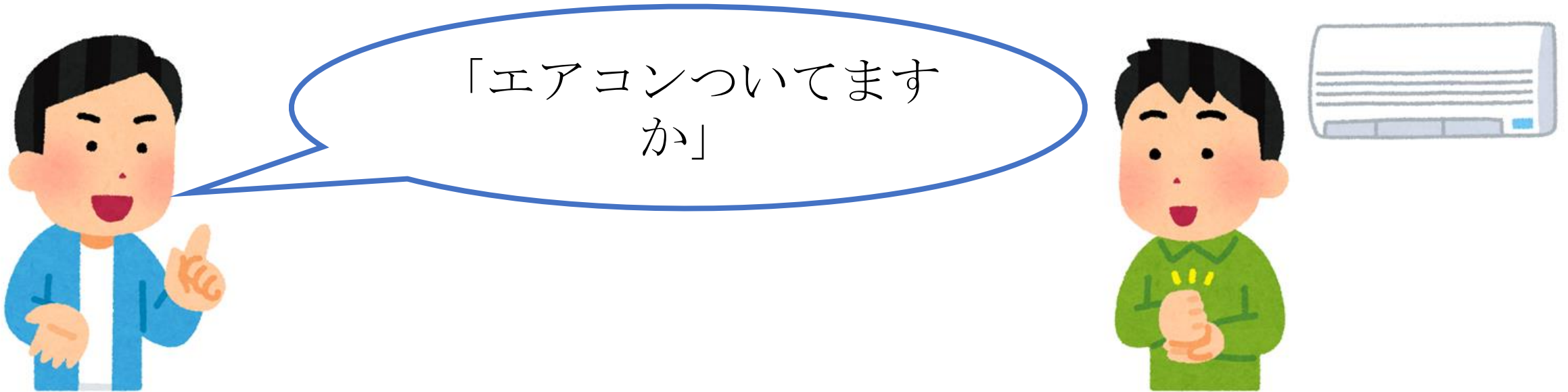
データサイエンス入門B #6

実例紹介5

音声処理におけるデータサイエンス・AIの活用

声で伝わる情報

30度以上ありそうな蒸し暑い部屋



声で伝わる情報

/e a k o N t s u i t e m a s u k a/

エアコン/ついて/ます/か

エアコンついてますか

エアコンという機械が動いていることを確認したい

エアコンをつけてほしい

音韻

単語

文

意味

意図



声で伝わる情報

誰(Aさん?)

年齢(20代?)

性別(男性?)

気分(いらだっている?)

態度(依頼している)

言語(日本語)

エアコンついてますか



音声とは

Aさん/年齢/性別/気分/態度/...

パラ言語情報

/e a k o N t s u i t e m a s u k a/

音韻

エアコン/ついて/ます/か

単語

エアコンついてますか

文

エアコンという機械が動いていることを確認したい

意味

エアコンをつけてほしい

意図

音声
=(伝統的には)
ことばを話す声

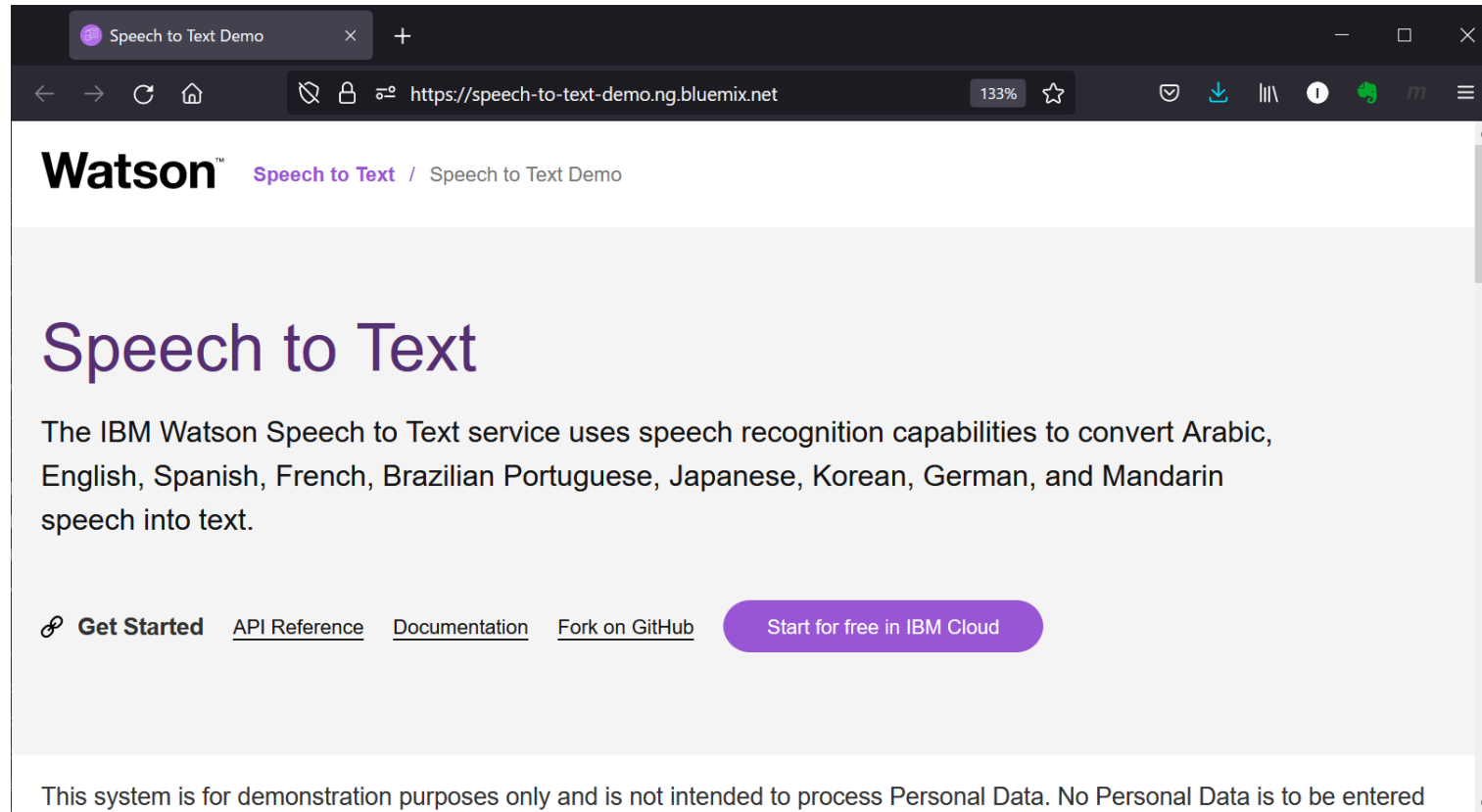
音声に関する技術

- レコーダ、プレーヤ、携帯電話
- 音声認識
- 音声合成、ボコーロイド
- 音声対話システム(音声アシスタント、スマートスピーカ)
- 音声強調

音声認識デモ

Watson speech to text API デモ


<https://speech-to-text-demo.ng.bluemix.net/>



音声合成デモ

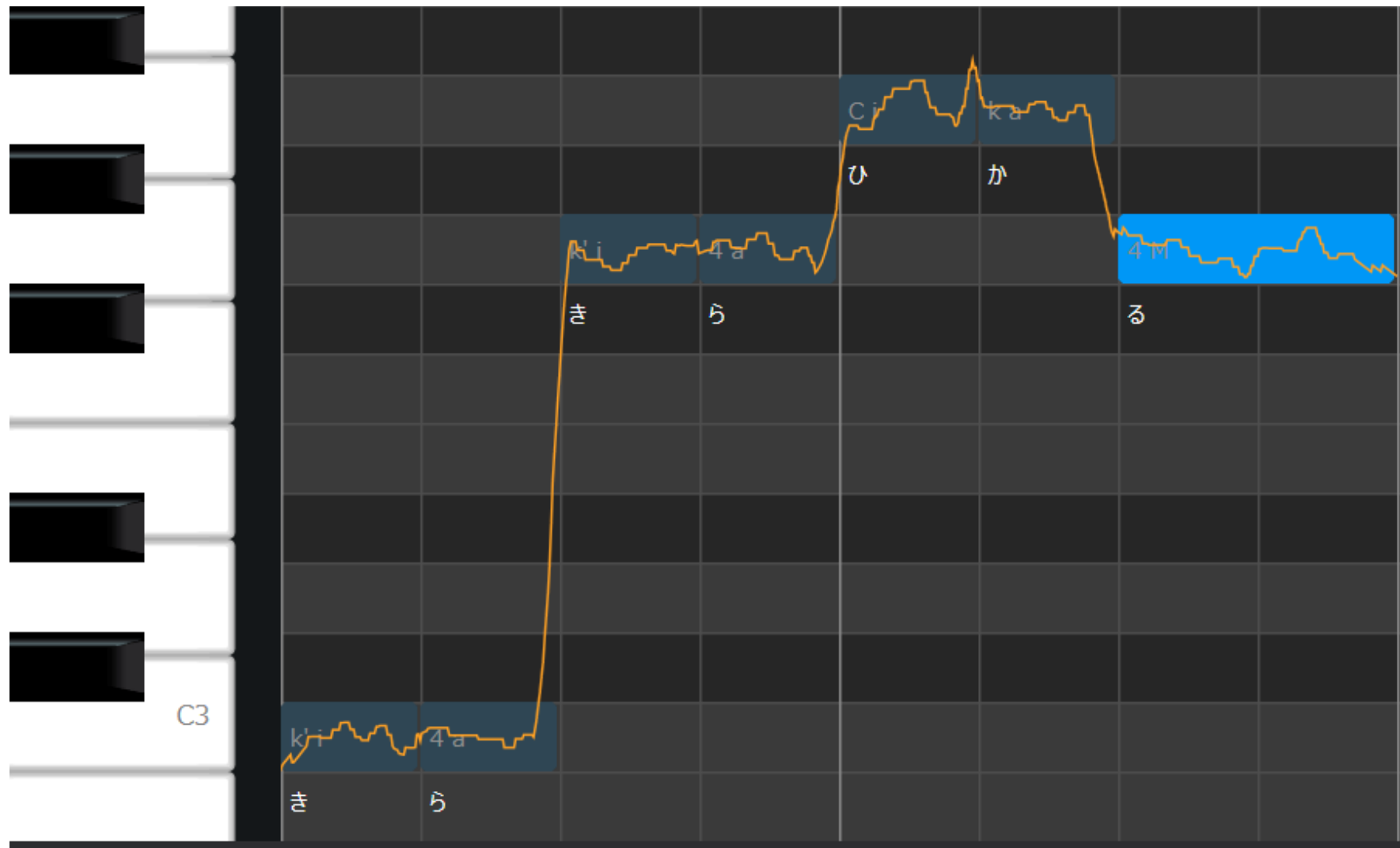
フェアリーデバイスズ mimi API 音声合成 TTS

<https://console.mimi.fd.ai/console/v1/speechsynthesis>



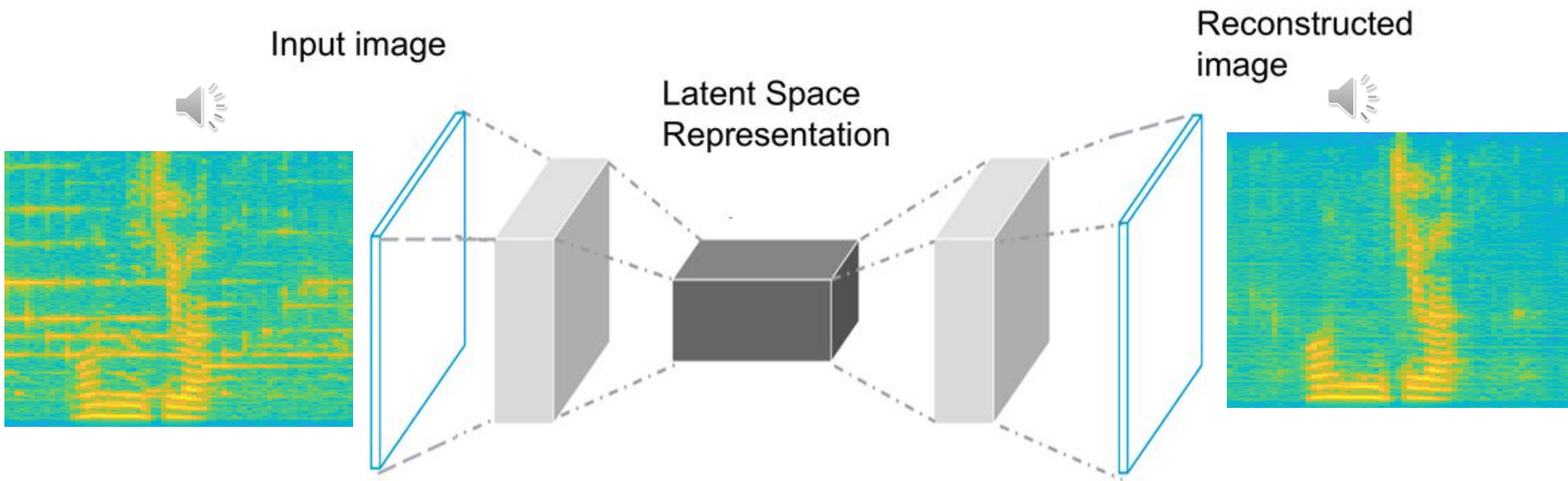
The screenshot shows a web browser window with the address bar displaying `https://console.mimi.fd.ai/console/v1/speechsynthesis`. The page title is "mimi® API Console". The main content area is titled "音声合成" (Speech Synthesis). It features two dropdown menus: "言語 *" (Language) set to "日本語" (Japanese) and "性別 *" (Gender) set to "女性" (Female). Below these is a text input field with the placeholder text "合成するテキストを入力してください*" (Please enter the text to be synthesized*). The browser's address bar shows a zoom level of 170%.

ボーカロイド



音声強調

U-Net, a Deep Convolutional Autoencoder



まとめ

- 音声認識
- 音声合成、ボコーロイド
- 音声強調