

MISHIMA.SYK#17 ONLINE 2021.06.12

音声認識で機器予約

@tkochi0603

自己紹介：@TKOCHI0603

▶ 伊豆半島にある製薬会社で、ITインフラ周りを担当してます

▶ 庭の芝生が育つように、土に穴を開けてる様子

▶ 土を踏むとザクっという音がいい

▶ 音声認識とは関係ありません



今日の話

- ▶ 前回#16ではChromeでの音声認識を紹介した
 - ▶ <https://github.com/kochi0603/mishima16>
- ▶ 小型デバイスを実験室に持ち込んで、音声認識を使ってみたい
 - ▶ 分析機器の予約で、使えるかしら？
 - ▶ 完全にシーズ駆動です。。。

Mishima.syk #16 2020/11/28

「気軽に音声認識」 @tkochi0603

- マイクを入力デバイスとして使えるか、音声認識を試した話。

環境

- Speech Recognition APIをChromeで試しました。(他のブラウザはまだまだ未実装みたい)
 - オンラインじゃないと動作しない ⇒ サーバに処理を丸投げしてるようです。
 - 実験的なAPIのため、未実装な機能もあるようです。

要件

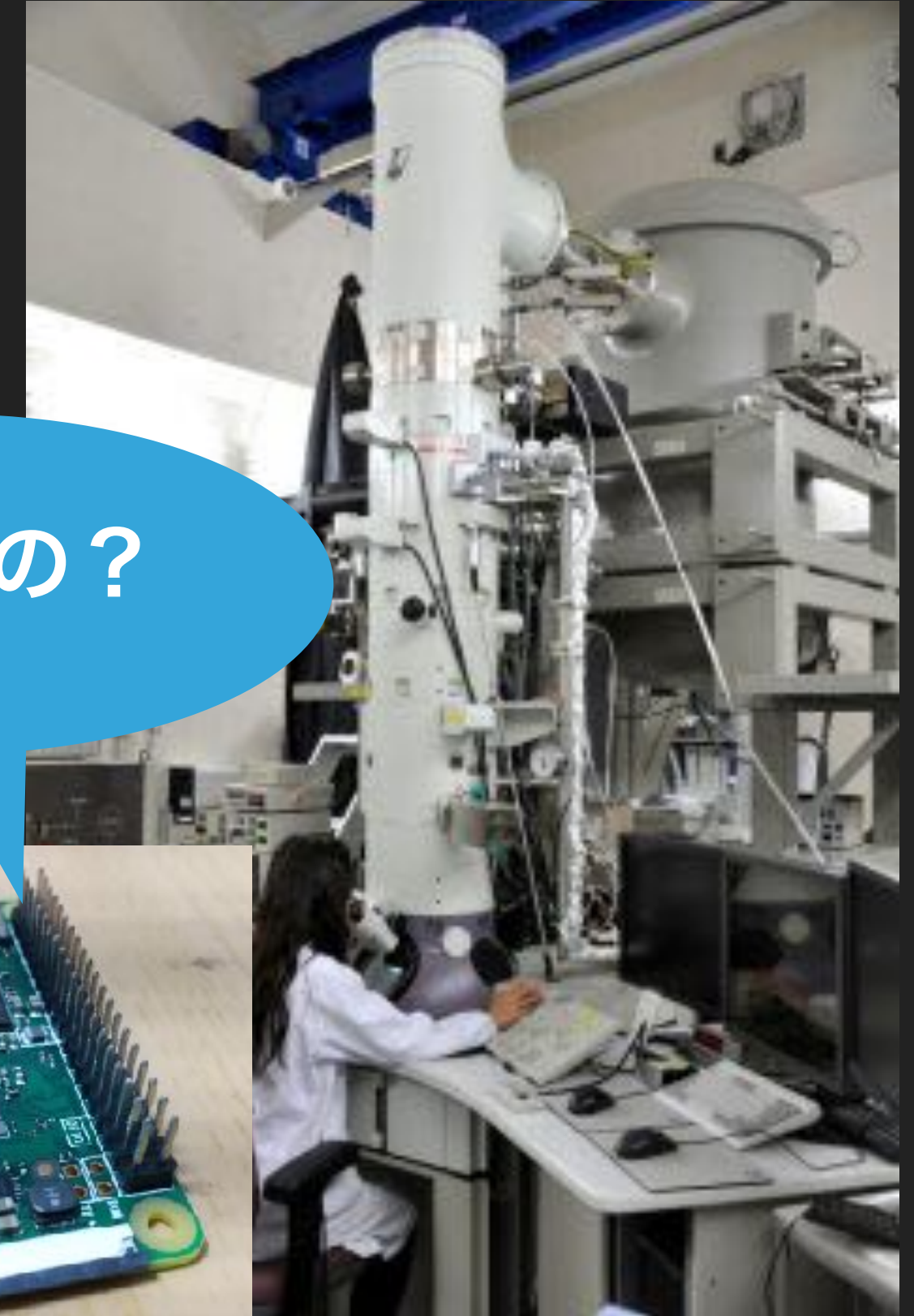
- ▶ マイク：音声認識
- ▶ スピーカーもしくはモニタ：
 - ▶ 返事するため。音声or表示
- ▶ ネットワークに接続できる
- ▶ サクッとできそうか
- ▶ 実験室に設置しやすい大きさ



使っていい？



使えんの？



引用 https://xlab.leica-microsystems.com/blog/learn/cryotem_nobel2017

流れ

- ▶ 予約したい分析機器の前に行って、そばにあるデバイスに、
 - ▶ ラムくん、空いてる？
 - ▶ 今予約中です
 - ▶ 2時間後まで空いてます
 - ▶ ラムくん、予約して
 - ▶ できません
 - ▶ 今から1時間予約しました

彼がラムくん



「誰」の予約なのかは
今回断念

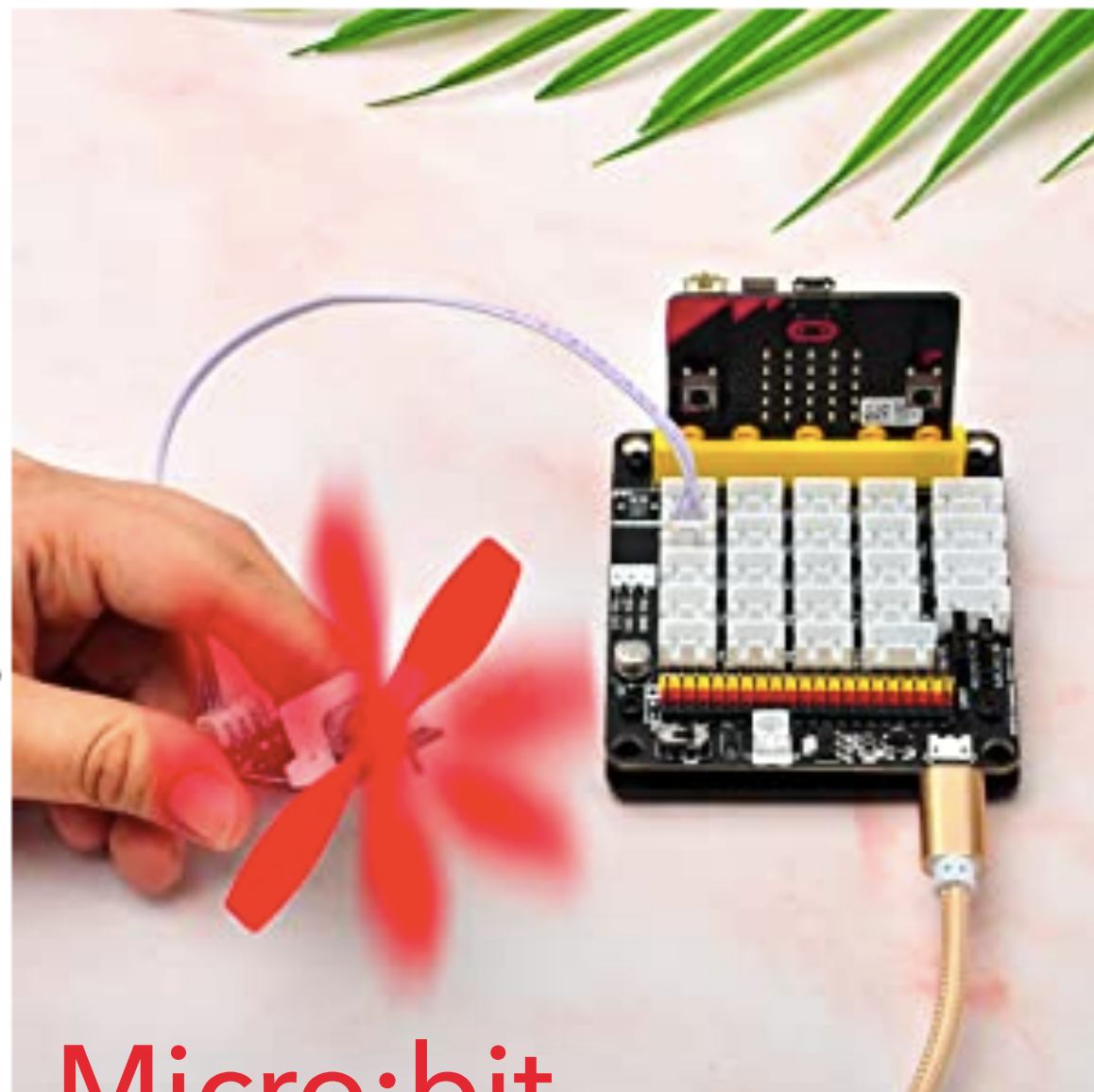
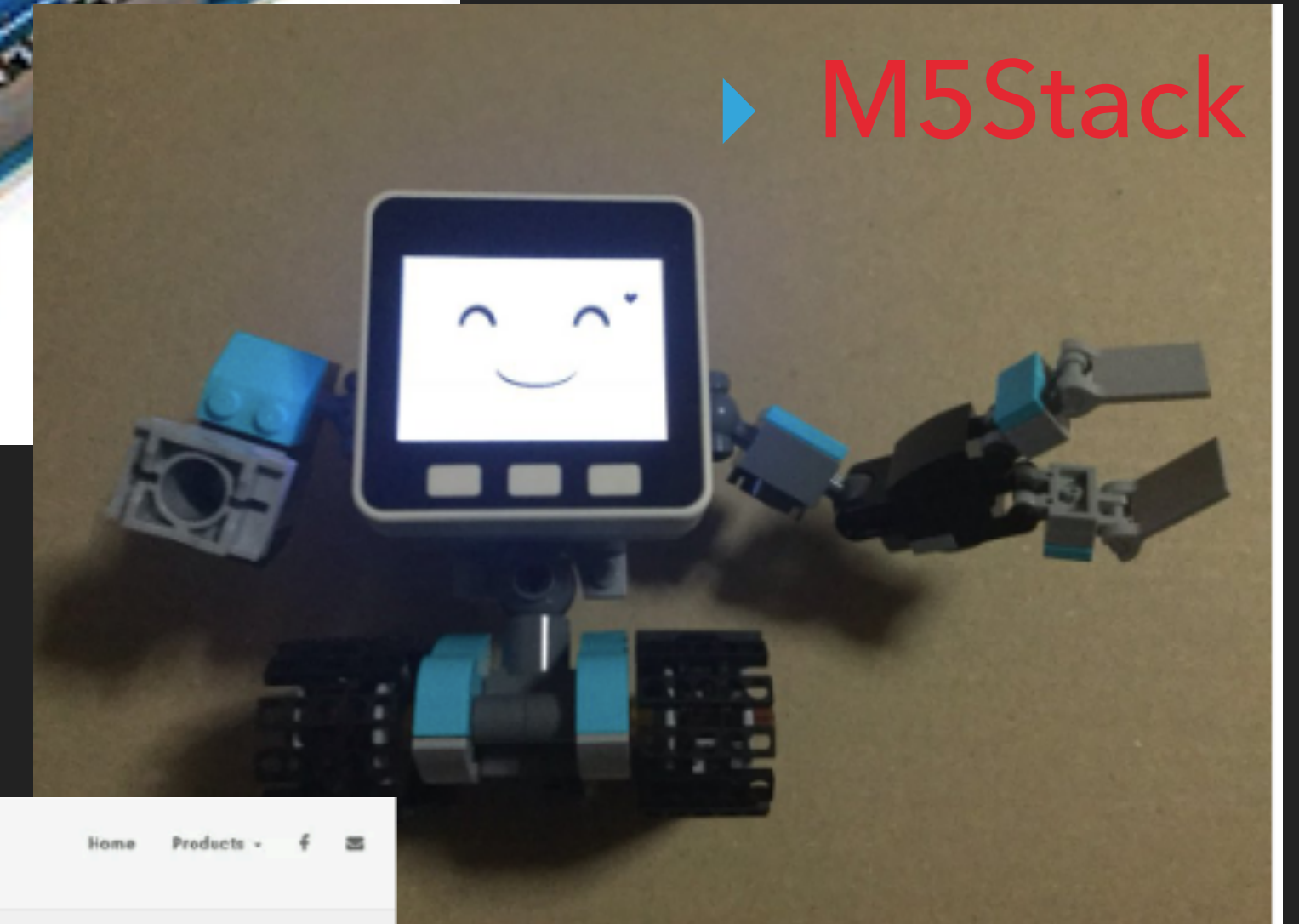
デバイス選び

- ▶ Arduino ? Micro:bit ? M5Stack ?
Raspberry PI ?

▶ Arduino



▶ M5Stack



▶ Micro:bit



▶ Raspberry PI

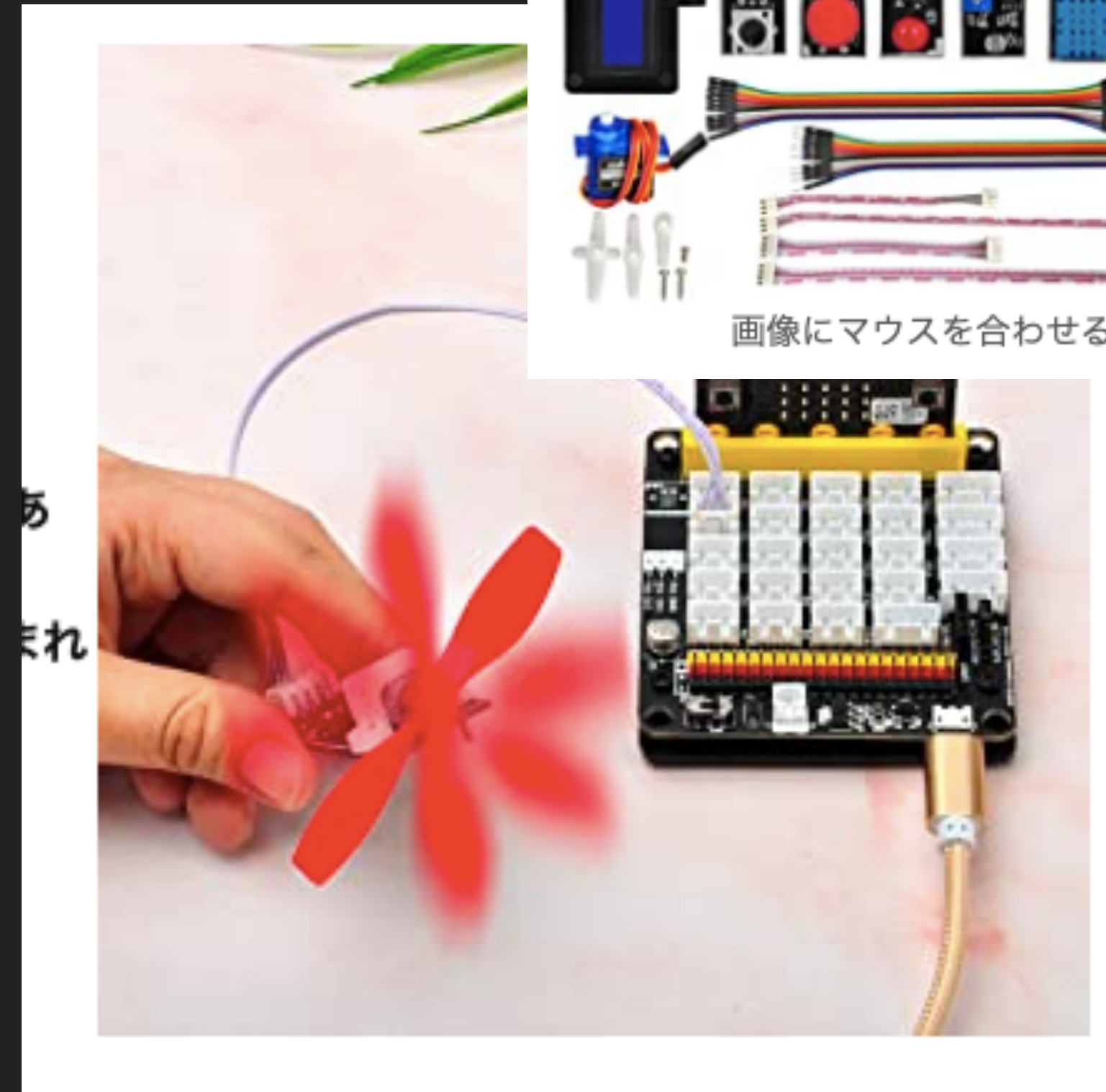
ARDUINO

- ▶ 知り合いから、センサー類に興味があるなら、Arduinoって聞いてた。
- ▶ C++っぽい言語
- ▶ 電子工作なArduinoに進むか、パソコンなRaspberry PIに進むか、迷うよねー



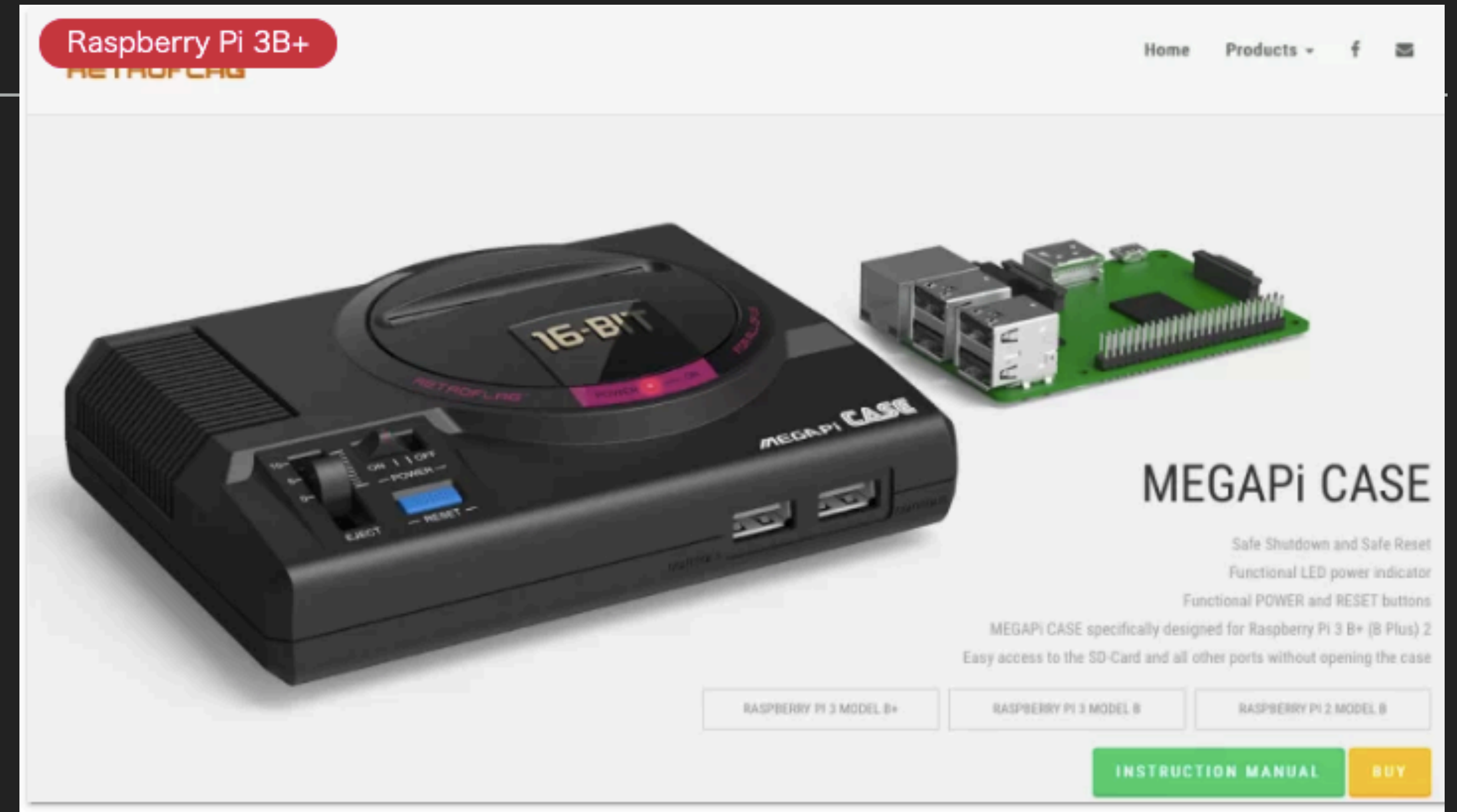
MICRO:BIT

- ▶ micro:bit は、学習材料として使われているので、子供と遊ぶと楽しそう。
- ▶ センサーの抜き差しが優しそう。
- ▶ 右図のセットだと5,000円越えだけど、必要な機能だけに限定すれば、お安くなるようだ



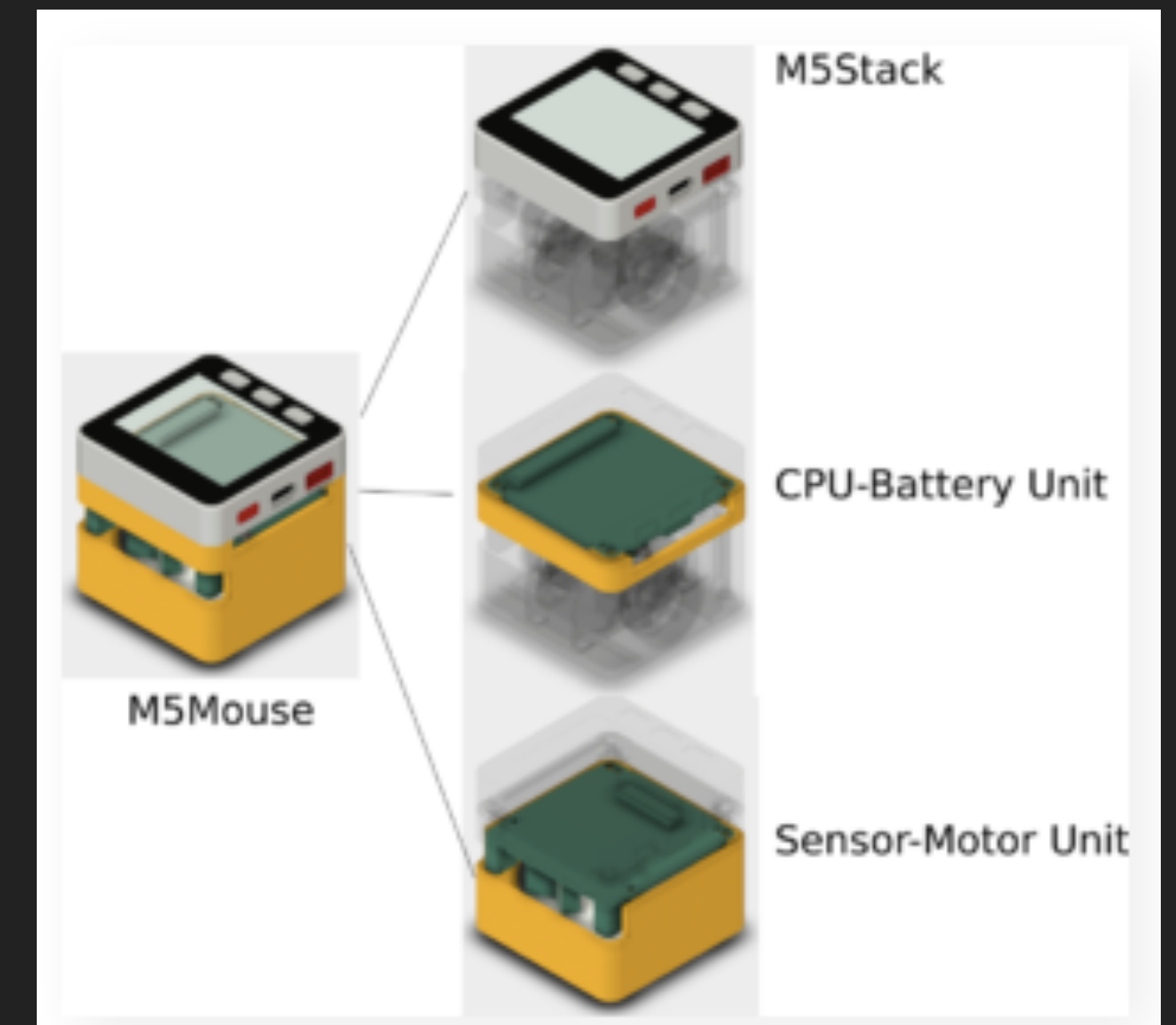
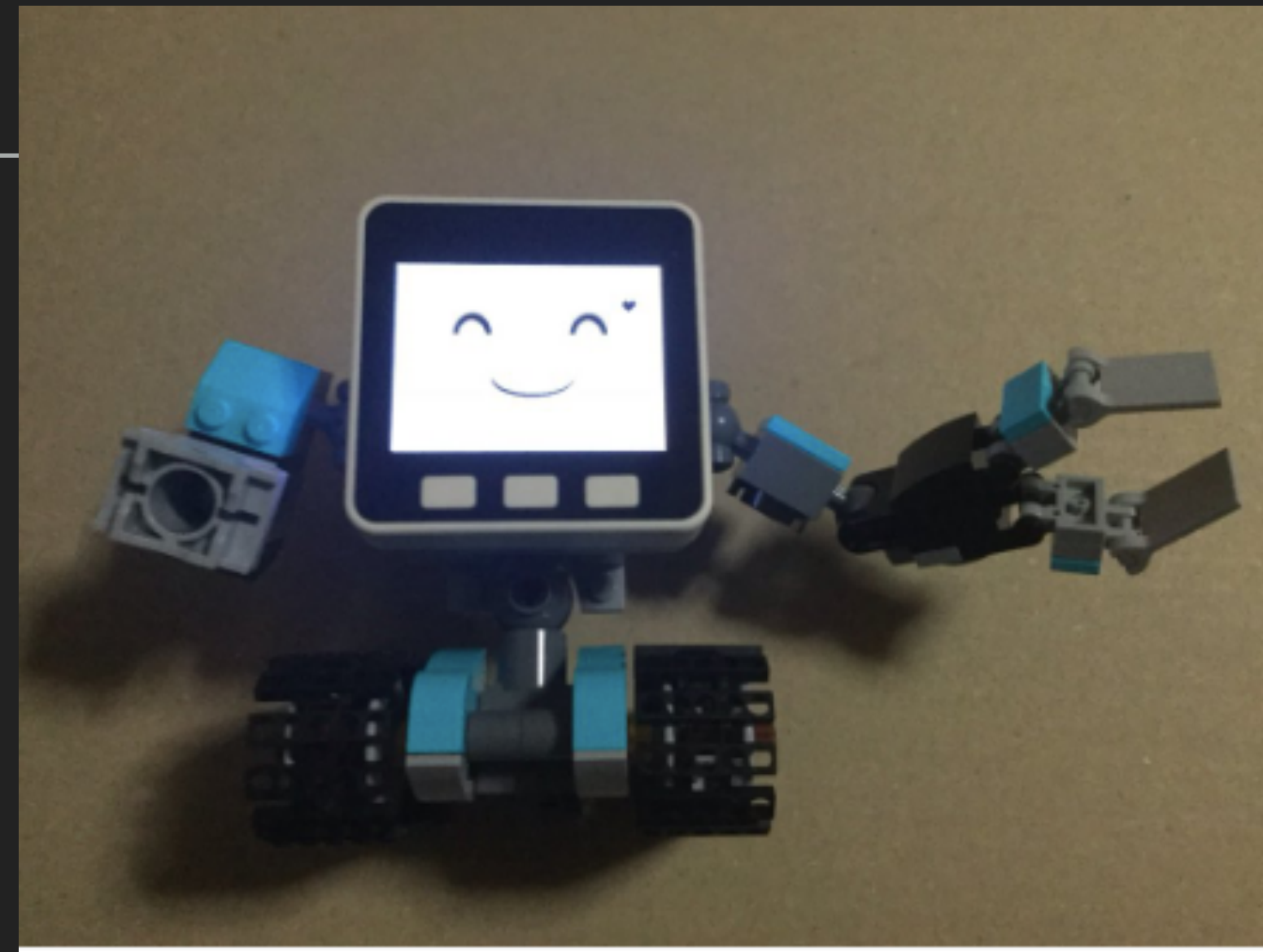
RASPBERRY PI 3B+ メガドラ

- ▶ いろんなケースがあるねー
- ▶ 実験室には置けないなー
- ▶ Raspberry pi 3B+ = 4,000円ぐらい
- ▶ ケース = 3,000円
- ▶ たっけーな



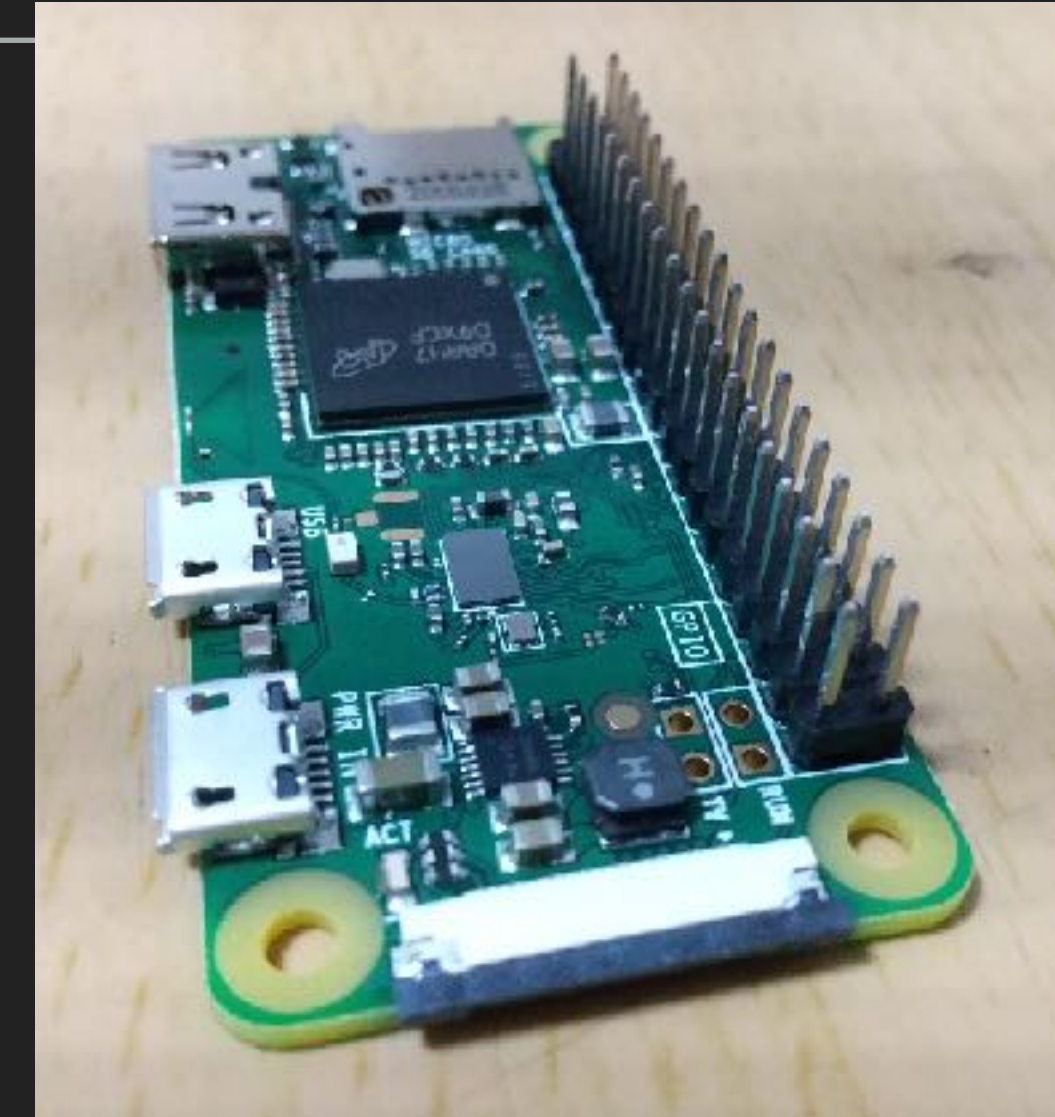
M5STACK

- ▶ モニタが最初から付いている
- ▶ 重ねるのが特徴
- ▶ レゴと組み合わせる
- ▶ とても惹かれたけど、結構値段が高い
- ▶ 本体で5,500円スタート



結局、RASPBerry PI ZERO WHにした

- ▶ ネットワーク周りに手こずるのが嫌なので、UbuntuなRaspberry PIに。ビビリです
- ▶ Raspberry PI zero WH(1,800円)を買った
 - ▶ 半田付けしてないzero W(1,300円)でもよかったな
- ▶ ケースは1,600円。奮発してしまった。
- ▶ あとは、micro USB(16GB) 800円ぐらい
- ▶ mini HDMIは必須ではない



BLUETOOTH
使わないと。

結局
7000円超え

手持ちのツールだけだと、全然スマートにならん。。

電源

USB変換

USBスピー
カー

USBマイク

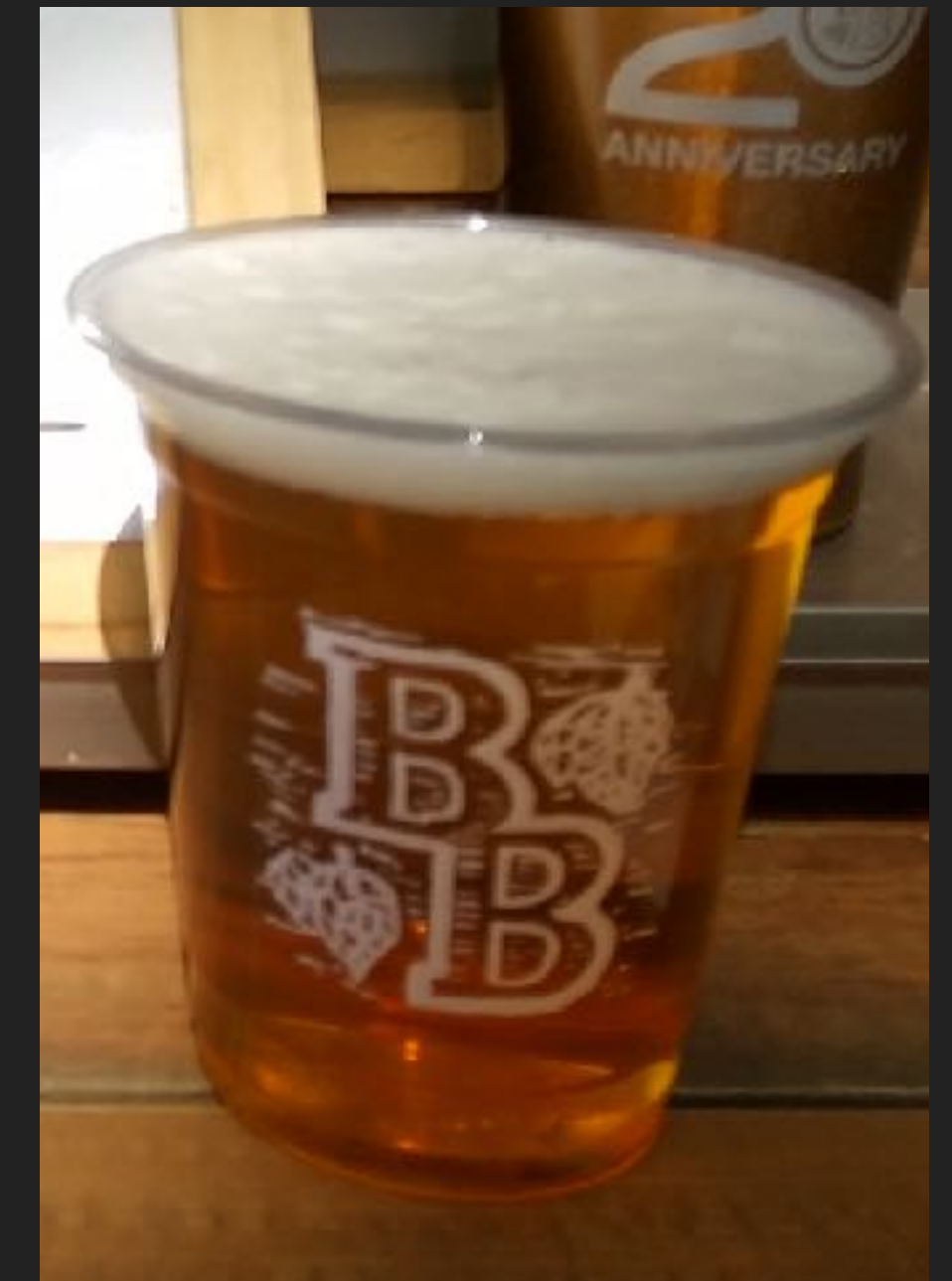
USB **2** のハブ



設定するなら、GUIありのOSイメージにするほうが良い

セットアップ

- ▶ GUIなしでインストールする方法
 - ▶ OSイメージをmicro SDに書き込んで、wifiの設定を行う
 - ▶ 起動したら、sshで接続するだけ
 - ▶ HDMIで出力できる方が安心ではあると思うけど。
 - ▶ あとは、マイクとスピーカーをラズパイに認識させる(USBで)



音声認識：JULIUS

- ▶ Juliusを試した（ラズパイでの事例も多い）
- ▶ 国産、オープンソース、オフラインで利用可
- ▶ 話し言葉の認識は精度低くて遅く、使えない
- ▶ 命令文(=アレクサ テレビ 付けて)は、精度高くて早い
 - ▶ 事前に設定しておいた辞書に存在する判定するだけ
 - ▶ 辞書を用意する必要がある
- ▶ なお、音声⇒文字列の認識なので、誰が喋ったかを判別するツールではない



音声認識：JULIUS

▶ どんな辞書？(一部)

- ▶ 同梱されたスクリプトが、ある程度やってくれるが、スクリプトを修正しないと動作しない

ラムくん

予約して

文法

```
[pi@raspberrypi:~/speechbot $ cat ../julius/dict/bot.grammar
S : NS_B TARGET ACTION NS_E
```

```
[pi@raspberrypi:~/speechbot $ cat ../julius/dict/bot.yomi
ラムくん らむくん
空いてる あいてる
予約して よやくして
```

表示したい文字と、読み方

```
[pi@raspberrypi:~/speechbot $ cat ../julius/dict/bot.voca
% TARGET
ラムくん      r a m u k u N
% ACTION
空いてる      a i t e r u
予約して      y o y a k u s h i t e
% NS_B
[s]           silB
% NS_E
[/s]          silE
```

表示したい文字列と、音素

音声合成

▶ OpenJTalk(国産)

▶ apt install でインストールできちゃう

▶ 文字列を入力すると、wavファイルが生成される

▶ 例

```
echo abc | open_jtalk -m /usr/share/hts-voice/  
mei/mei_happy.htsvoice -x /var/lib/mecab/dic/  
open-jtalk/naist-jdic/ -ow 1.wav
```

男性の声か、
女性の声か

文節
解釈

出力
ファイル

Open JTalk



What is Open JTalk?

Open JTalk is a Japanese text-to-speech system.
This software is released under the [Modified BSD license](#).

Getting Open JTalk

Open JTalk version 1.11 (25 December, 2018) **NEW!**

The Japanese text-to-speech system.

It works in Linux machines.

[Documentation](#)

[Source Code](#)

やりたいことが書いてあるページがあった

- ▶ ぜひ見てください。
- ▶ https://zenn.dev/minako_ph/articles/45503855f1f626

結局真似した





RaspberryPi4で音声認識ボットを作るハンズオン【ラズパイ+Julius+Open JTalk】

📅 2021.05.09に公開 🔄 2021.05.10に更新 ⌚ 13 min read

こんにちは、ニホンザルとネコが大好きな @minako-ph です 🐱

最近こんなを作ったので、忘れないうちにこの時の知見をまとめつつ
皆さんにも気軽にラズパイに挑戦してもらえるようハンズオン形式でまとめてみます 🐱!

Topics

- # 電子工作
- # ラズパイ
- # julius
- # openjtalk
- # 音声認識
- # Tech

PYTHON

- ▶ Pythonでコードを書いた
 - ▶ 音声認識をした命令文を判断し、
 - ▶ 予約表のテキストファイルを確認して、
 - ▶ 音声合成で返事する

できた

まとめ

- ▶ 音声認識で、命令文にすると、「誰が」や「いつ」を伝えるのが難しい = 予約表では必須
 - ▶ 組織内に設置するなら、あらかじめ全員の名前や、時間のパターンを辞書にする？
 - ▶ 「誰が」「いつ」「何を」「どうする」の、どれかを割愛できると良い
 - ▶ 分析機器1台ごとに予約デバイスを用意すれば、「何を」を割愛できる
 - ▶ 実験者1名ごとに予約デバイスを用意すれば、「誰が」を割愛できる
 - ▶ など考える
- ▶ 音声合成はいいですね。遊べそうです。
- ▶ もっとスマートなデバイスにしないと実験室には置けない。。