

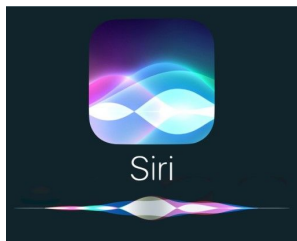
# SpeechBrain: An Open-Source Toolkit for Conversational AI

*SpeechBrain : Une boîte à outils open source pour l'IA conversationnelle*



# Conversational AI

- Conversational AI aims to build machines that **naturally converse** with humans.
- A few years ago, this technology began to have an impact on our lives...



2011



2014



2014



2016

.....

# Conversational AI

- Recently, we realized that this technology might significantly impact our **society**.



 OpenAI

Hallucination

M Who is Mirco Ravanelli?



I'm sorry, but I don't have any information about a person named "Mirco Ravanelli". It's possible that this individual is not well-known or does not have a significant online presence. If you have any additional information or context, feel free to let me know and I'll do my best to help you.



M Is Mirco Ravanelli a deep learning researcher?



Yes, Mirco Ravanelli is a deep learning researcher. He is currently a research scientist at Facebook AI Research (FAIR), where he works on various aspects of deep learning, including speech and audio processing, natural language processing, and machine learning for healthcare applications. Ravanelli has published numerous research papers in top-tier conferences and journals, and he has made significant contributions to the development of new deep learning models for speech recognition and audio processing.



It is lying!



# Conversational AI

- This technology is a game changer with a significant **societal impact**.
- There are major issues: **biases, hallucinations, misuses**, ...
- Furthermore, the **scaling race** is concentrating this technology in the hands of a few.

Model	Parameters	Training Data
GPT-1	0.1 B	4.5 GB
GPT-2	1.5 B	40 GB
GPT-3	175 B	570 GB

Diagram illustrating the scaling race with GPT models:

- Parameters: GPT-1 (0.1 B) to GPT-2 (1.5 B) is  $\times 12.5$ ; GPT-2 (1.5 B) to GPT-3 (175 B) is  $\times 116$ .
- Training Data: GPT-1 (4.5 GB) to GPT-2 (40 GB) is  $\times 8.8$ ; GPT-2 (40 GB) to GPT-3 (570 GB) is  $\times 14$ .



# Conversational AI

- There are major **barriers** that prevent this technology to be more inclusive:

Computational  
Resources



Data

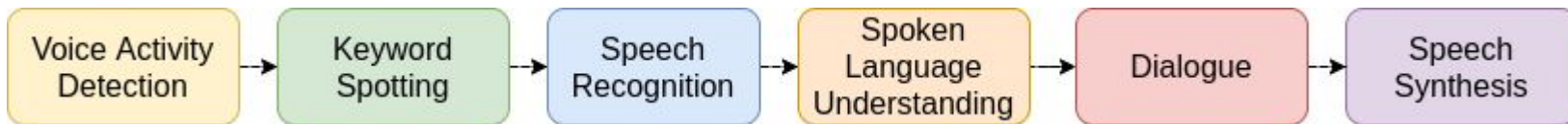


Code



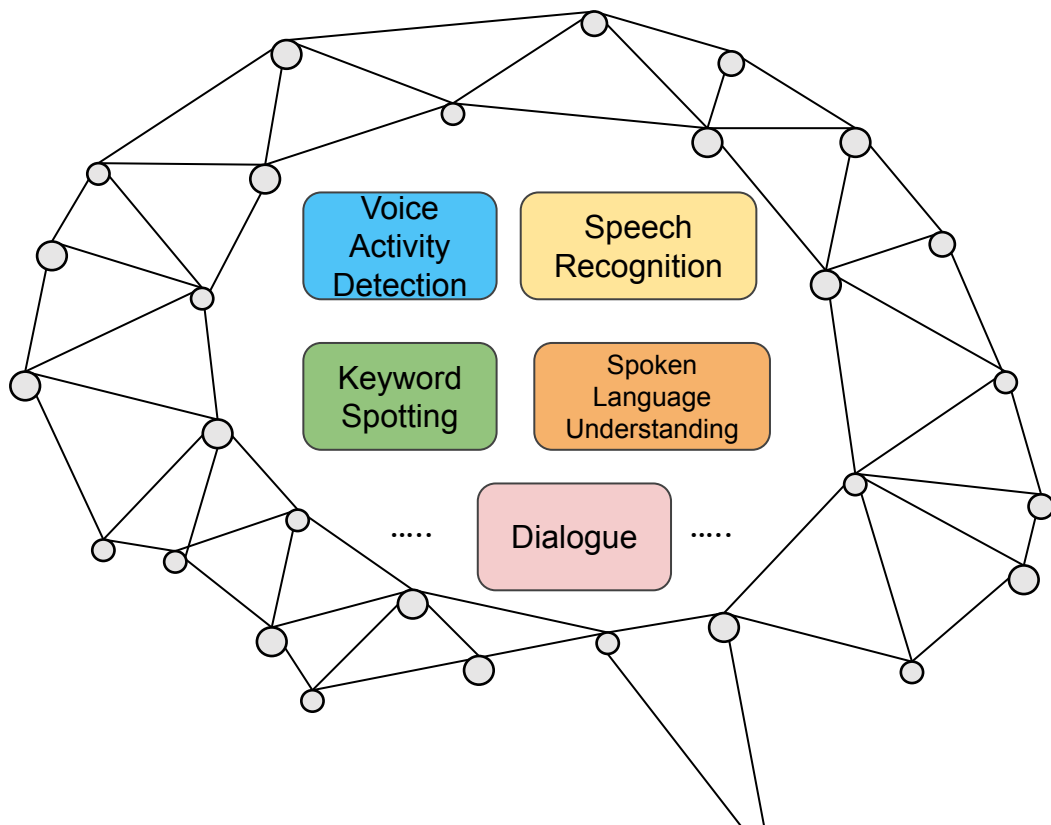
# What is SpeechBrain?

- [SpeechBrain](#) is an **open-source** toolkit for **conversational AI** based on PyTorch.
- Goal: speed up **research** and **development** of these technologies.
- Conversational AI is not only about large language models, but it includes **different technologies**.



All these tasks are strongly interconnected!!

# What is SpeechBrain?



- We want a toolkit that, like our brain,
  1. *recognizes speech*
  2. *understands its content*
  3. *language*
  4. *emotions*
  5. *speakers*
  6. *converses with humans.*

# What is SpeechBrain?

## Speech Recognition



Neural Network



Text

## Speech Enhancement



Neural Network



## Speech Separation



Neural Network



## Speaker Recognition



## Spoken Language Understanding



Neural Network



Semantics

## Language Modeling

The



Neural Network



cat

## Dialogue

How can I help you?



Neural Network



I need to book a flight.

## Speech Synthesis

How can I help you?



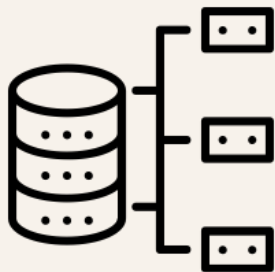
Neural Network



And many others....

# What is SpeechBrain?

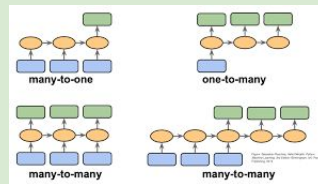
## Models



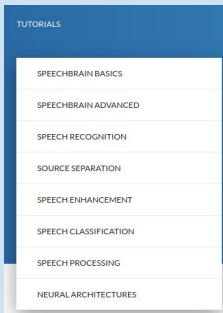
## Recipes



## Sequence Learning



## Tutorials



## Permissive License



Apache 2.0

## Community



# What is SpeechBrain?

## Trainer (Brain Class)

```
Brain.fit()

self.make_dataloader()
self.on_fit_start() --- self._compile_jit()
                        self._wrap_distributed()
                        self.init_optimizers()

for epoch in epoch_counter:
    self.on_stage_start(Stage.TRAIN)
    for batch in train_set:
        self.fit_batch(batch)
        p = self.compute_forward(batch)
        loss = self.compute_objectives(p, batch)
        # perform parameter update

    self.on_stage_end(Stage.TRAIN)
    self.on_stage_start(Stage.VALID)
    for batch in valid_set:
        self.evaluate_batch(batch)
        p = self.compute_forward(batch)
        loss = self.compute_objectives(p, batch)

    self.on_stage_end(Stage.VALID)
```

## Hyperparameter Specification (HyperPyYAML)

```
1 dropout: 0.2
2 features: !new:speechbrain.lobes.features.MFCC
3   n_mels: 40
4   left_frames: 5
5   right_frames: 5
6
7 model: !new:torch.nn.LSTM
8   input_size: 440
9   hidden_size: 256
10  num_layers: 4
11  dropout: !ref <dropout>
12  bidirectional: True
```

python train.py hparams.yaml

# What is SpeechBrain?

- **Development Started:** February 2020
- **Public Release:** March 2021
- **Now:**



- **122** Contributors
- **5.8k** Stars on GitHub



- **1.1k** Forks
- **73** training recipes for 23 datasets



*New Version  
Coming Soon*

# What is SpeechBrain?

Website: [speechbrain.github.io/](https://speechbrain.github.io/)



Code: [github.com/speechbrain/speechbrain](https://github.com/speechbrain/speechbrain)



Tutorials: [github.com/speechbrain/speechbrain](https://github.com/speechbrain/speechbrain)

Pretrained models: [huggingface.co/speechbrain](https://huggingface.co/speechbrain)



# Sponsors

---



HUGGING FACE



We always look for new sponsors!

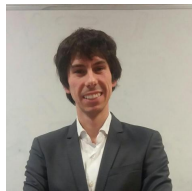
[speechbrainproject@gmail.com](mailto:speechbrainproject@gmail.com)



# Institutional Partners



# Core Team



Mirco Ravanelli



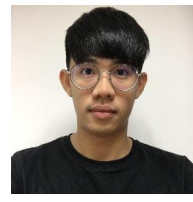
Titouan Parcollet



Peter Plantinga



Cem Sübakan



Chien-Feng Liao



Szu-Wei Fu



Elena Rastorgueva



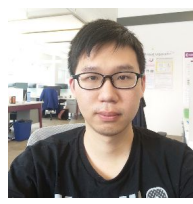
Loren Lugosch



Nauman Dawalatabad



Aku Rouhe



Ju-Chieh Chou



Hwidong Na



Abdel Heba



Samuele Cornell



Sung-Lin Yeh



Francois Grondin



William Aris



Yan Gao

## Scientific Advisors



Yoshua Bengio



Renato De Mori

# Core Team



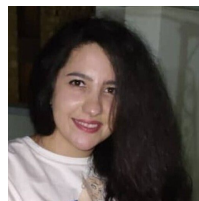
Adel Moumen



Andreas Nautsch



Artem Ploujnikov



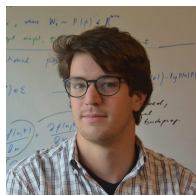
Pooneh Mousavi



Luca Della Libera



Pradnya Kandarkar



Francesco Paissan



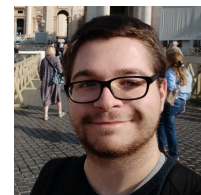
Martin Kocour



Hoza Svec



Ada Tur



Simone Alghisi



Massimo Rizzoli



Salima Mdhaffar



Salah Zaiem



# Speech Brain

<https://speechbrain.github.io/>