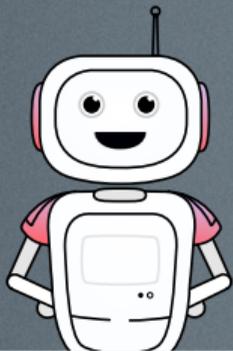


# Common Voice – DeepSpeech

Données et modèles pour la reconnaissance vocale libre



Alexandre Lissy  
alissy@mozilla.com



# Common Voice

## État & Objectifs

### Cassons les barrières de la voix

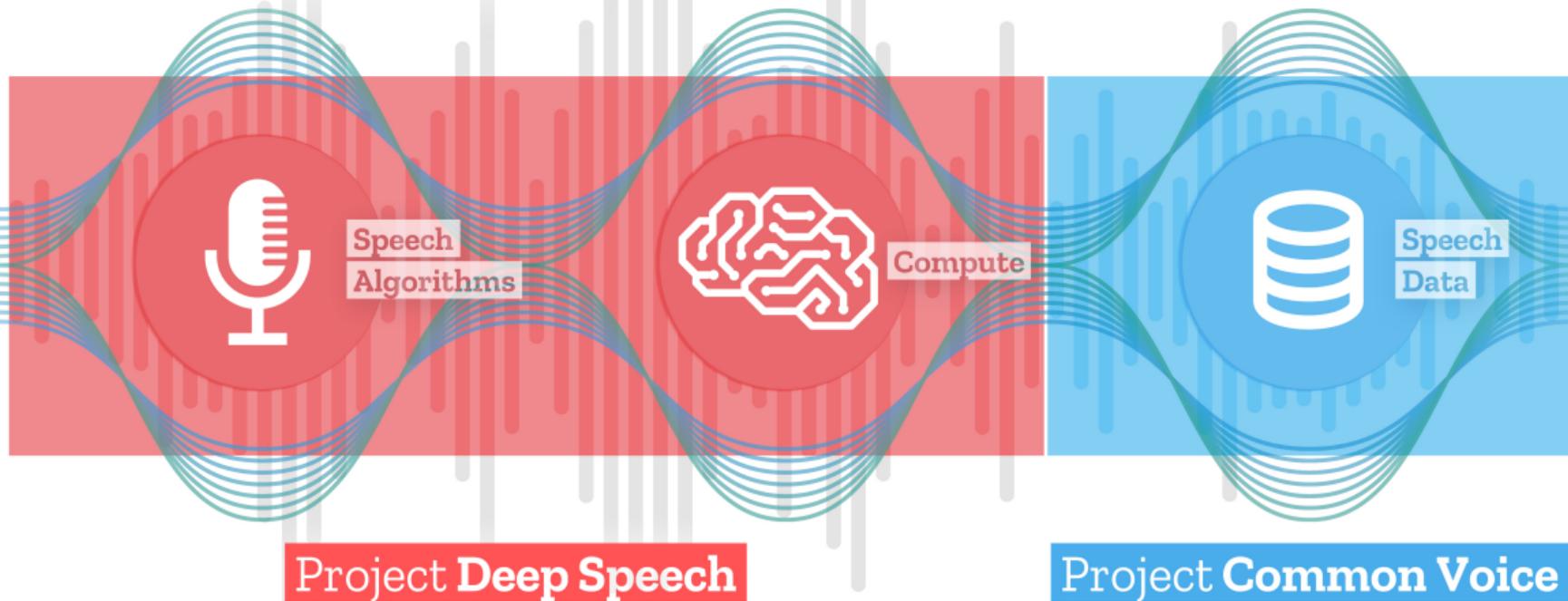
- Prolifération des systèmes contrôlables à la voix (objets connectés, etc.)
- Accès au Web de plus en plus via la voix
- Historique et facilité : concentration sur quelques services en ligne

### Cassons les barrières de la voix

- Accès : peu de gros jeux de données publiquement accessibles
- Coût : ticket d'entrée important pour les jeux de données commerciaux, avec des complexités légales
- Biais : seuls certains marchés sont bien pourvus

# moz://a Speech Recognition

## Ingredients



# Common Voice

## Objectifs

### Collecte de données ouvertes pour l'apprentissage machine

- Données : **audio** et **texte** correspondant
- Diversité
  - Langues
  - Accents
  - Ages
  - Genres
  - Qualité sonore

Nous visons 10 000 heures de données validées par langue :

<https://voice.mozilla.org>

# Accompagner la communauté

Mozilla ne serait rien sans sa communauté

## Contributions par les faits

- 38 langues actives, dont l'anglais, l'allemand, le kabyle, le français, le catalan, le breton, le basque.
- 1200 heures validées en novembre 2018 au total, 3300 heures validées en novembre 2019, 4000 heures enregistrées; contributeurs de 112 pays.
- Dont 324 heures enregistrées et 284 validées en français.
- Plus de 110000 contributeurs individuels ont enregistré et/ou validé.
- Plus de 200 développeurs ont participé.
- Mycroft et Snips (assistants numériques), Université de Bangor (Pays de Galles).
- Jeu de données téléchargé des centaines de fois tous les mois.

# DeepSpeech : proposition de valeur

## Outils, modèle et API

- Des outils pour entraîner et évaluer un modèle de reconnaissance vocale
- Un(des) modèle(s) de qualité production mis à disposition librement
- **Une API C** (avec des interfaces) (et la plus simple)

## Performances

- Temps réel sur desktop, Snapdragon 820/835, RPi4
- **v0.6.0** 7.5% *WER* en anglais

Déjà intégré (beta) dans eSup Pod v2.4 pour sous-titrage automatique en Français

## Aller plus loin

- Site officiel Common Voice <https://voice.mozilla.org>
- Dépôt Github Francophone Common Voice / DeepSpeech <https://github.com/Common-Voice/commonvoice-fr>
- Dépôt Github DeepSpeech <https://github.com/mozilla/DeepSpeech>
- Forum Discourse Common Voice Francophone <https://discourse.mozilla.org/c/voice/fr>
- Forum Discourse DeepSpeech <https://discourse.mozilla.org/c/deep-speech>

**moz://a**