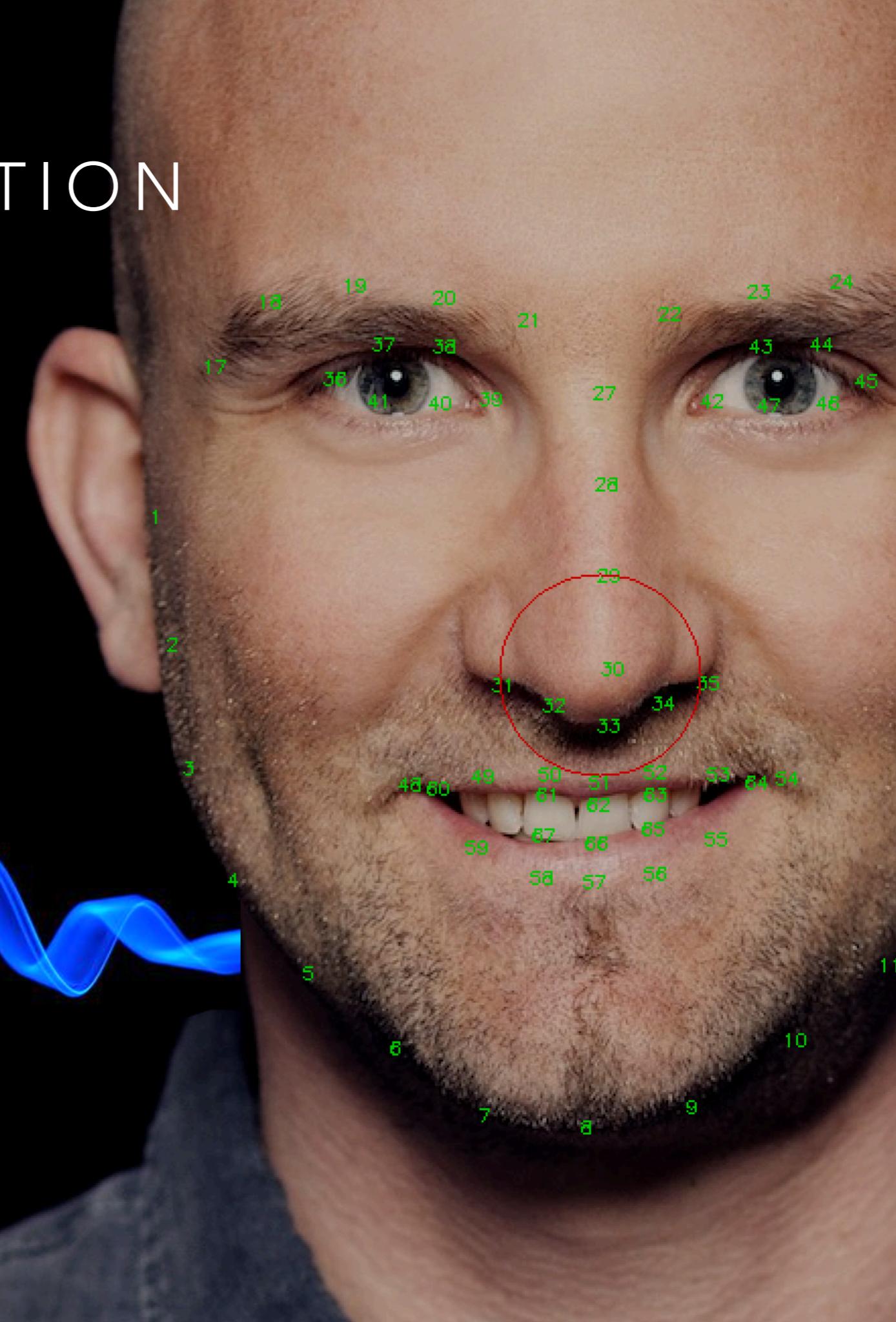
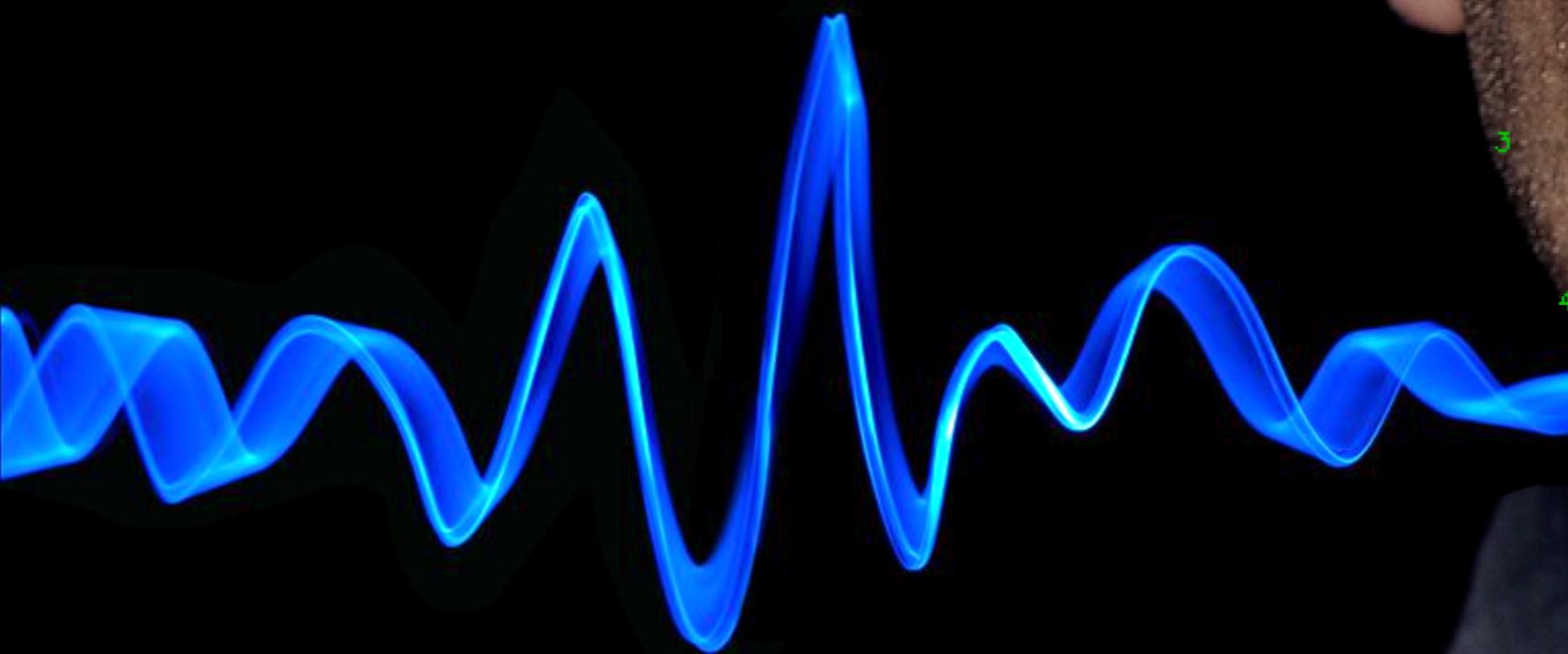


TELECOM PARISTECH - 2019

# MULTIMODAL EMOTION RECOGNITION



FIL ROUGE  
CONTEXTE ET  
ENJEUX



# CONTEXTE

- Pôle Emploi : Etablissement public chargé de l'emploi en France.
- Missions :
  - Faciliter le retour à l'emploi
  - Accompagner dans l'orientation
  - Appuyer le recrutement
  - Simplifier l'accès aux services



# CONTEXTE

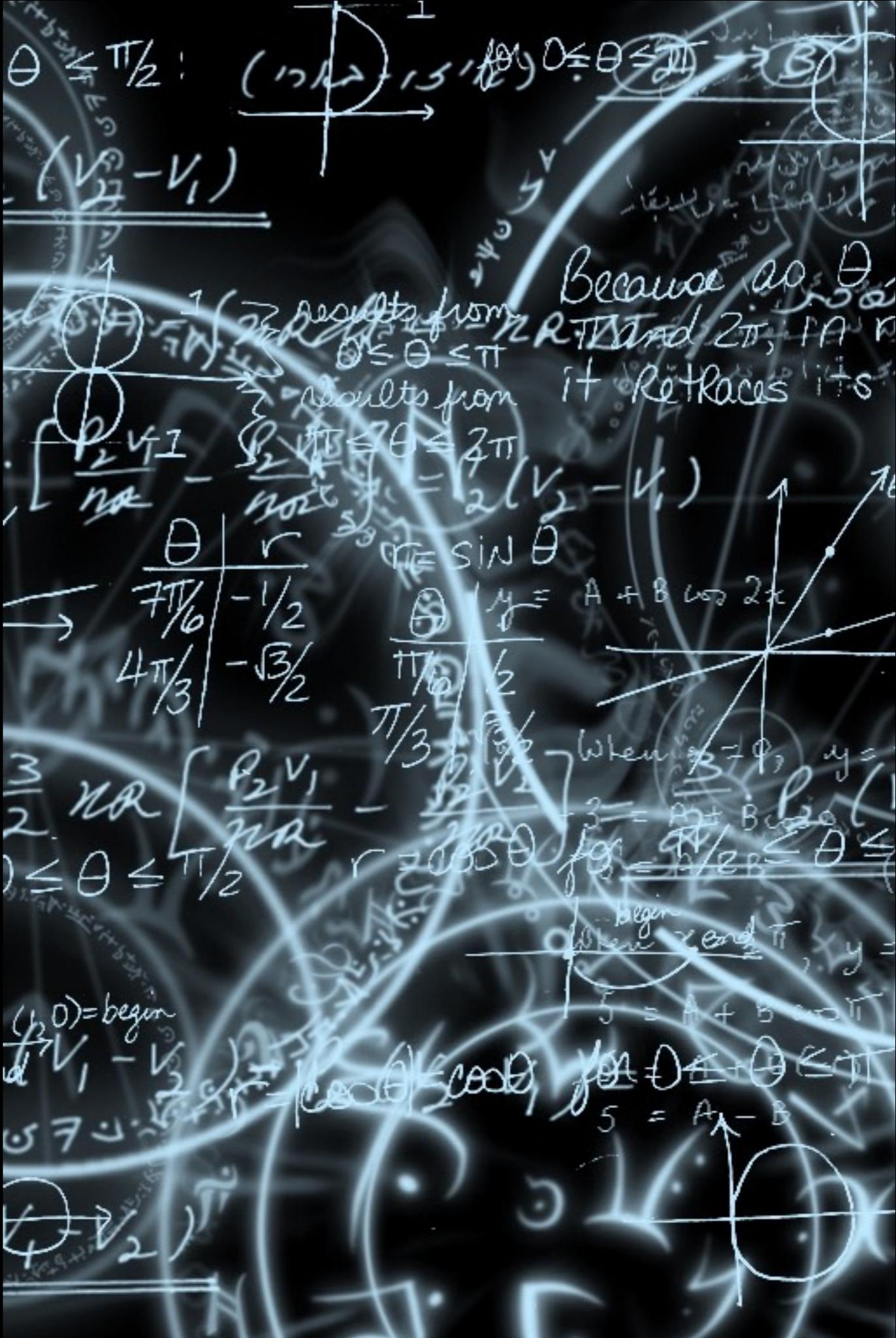
- Informatique affective : *Etude et développement de systèmes ayant les capacités de reconnaître, d'exprimer, de synthétiser et modéliser les émotions humaines.*
- Usages pour le relation client :
  - Assistance aux conseillers
  - Adapter et catégoriser les échanges
  - Suivi et personnalisation

# ENJEUX

- Tester et concevoir des solutions d'analyse des émotions :
  - Texte : Réponses à des questions en entretien
  - Voix : Enregistrements d'échanges avec un conseiller
  - Vidéo : Enregistrements d'échanges avec un conseiller
- A partir d'Open Data
- Pour élargir le spectre des services offerts

FIL ROUGE

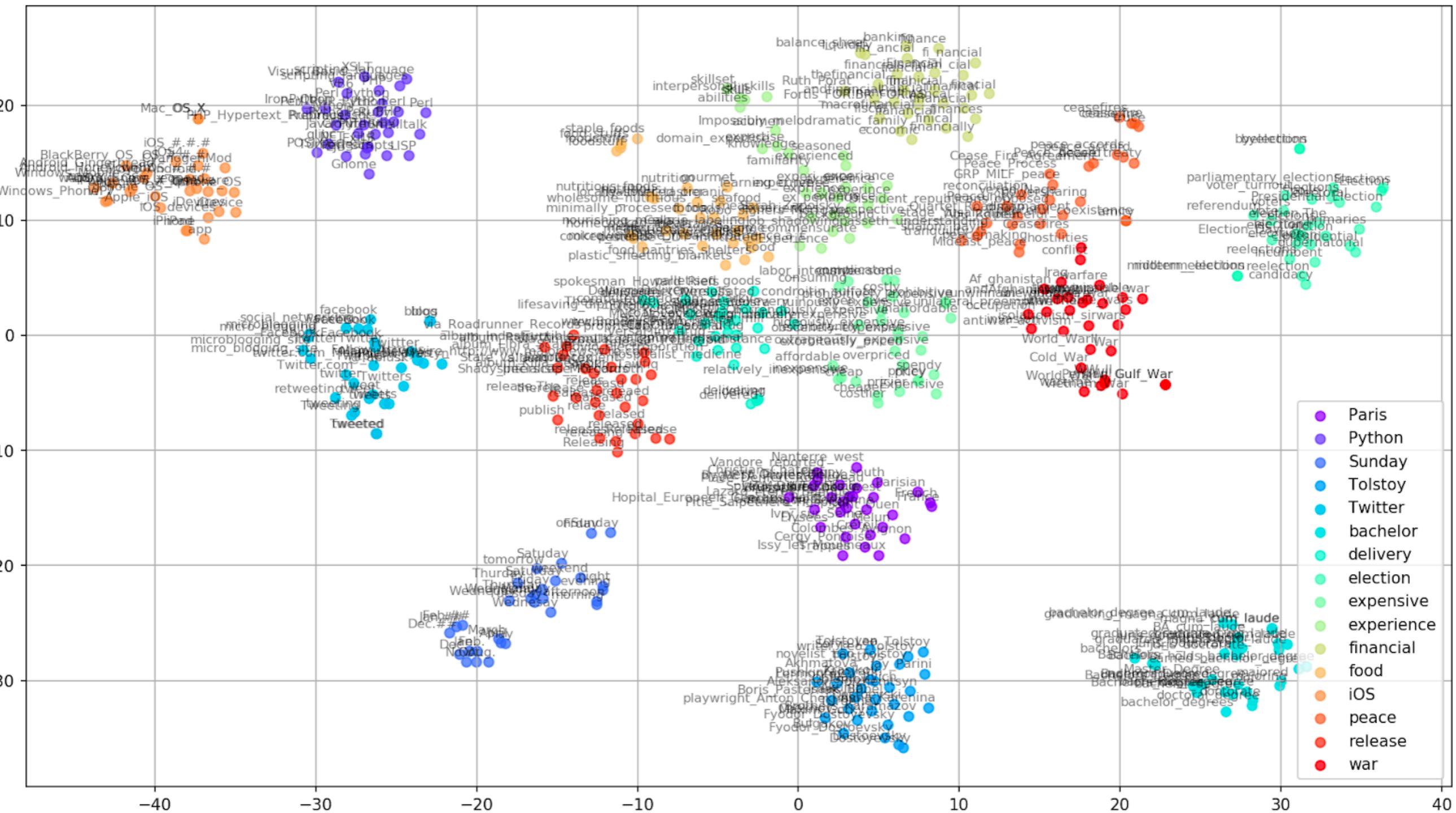
ETAT DE L'ART



# TEXTE

- Identifier des traits de personnalité (Big5) à partir de réponses à des questions spécifiques
- Natural Language Processing :
  - Pré-Traitement
  - Embedding
  - Classification
  - *Linguistique..*
- Dernières avancées :
  - Recurrent Neural Networks (LSTM, GRU)
  - Data Visualization (t-SNE, WordCloud, Topic Modelling)

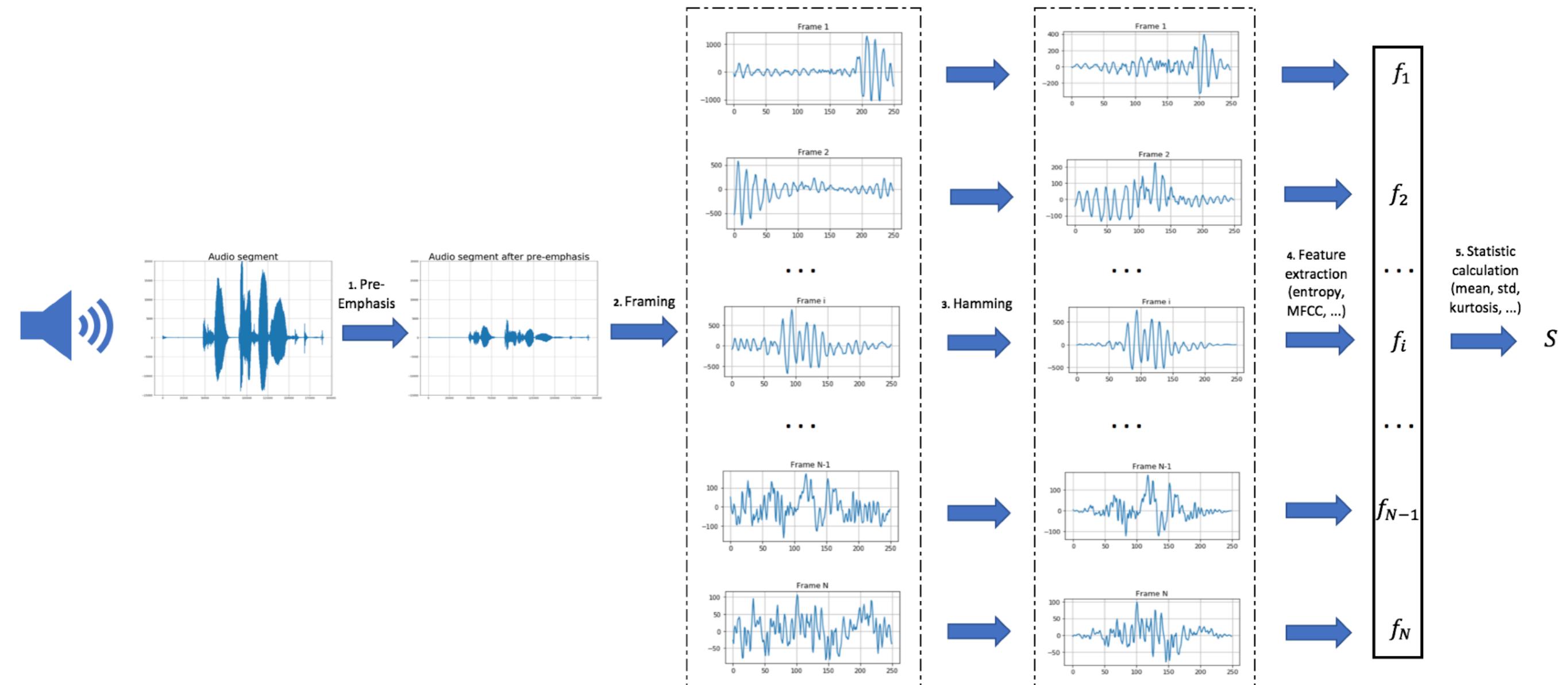
# T-SNE



# AUDIO

- Identifier des émotions à partir de la voix
- Signal Processing :
  - Pré-Traitement du signal : Pré-Emphasis, Framing, DFT
  - Extraction de variables explicatives (temps, fréquence)
  - Classification (SVM)
- Dernières avancées :
  - Classification : CNN, modèle hybride
  - Temps réel

# SIGNAL PROCESSING

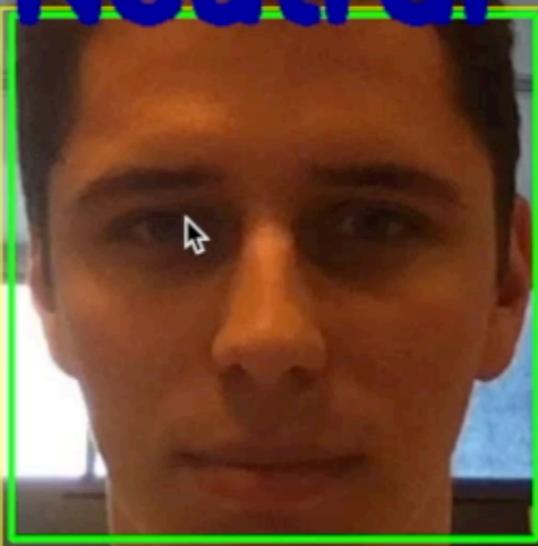


# VIDEO

- Identifier des émotions à partir d'une vidéo
- Image Processing :
  - Pré-Traitement de l'image : Scaling, Extraction du visage
  - Extraction de points sur le visage
  - Classification (SVM)
- Dernières avancées :
  - Classification : CNN, modèle hybride
  - Temps réel

Angry : 0.167  
Disgust : 0.088  
Fear : 0.039  
Happy : 0.218  
Sad : 0.164  
Surprise : 0.076  
Neutral : 0.248

Neutral



FIL ROUGE  
OBJECTIFS  
POURSUIVIS



# ACCURACY

Texte

**45%**



**58%**

Audio

**75%**



**85%**

Video

**61 %**



**75%**

# ENSEMBLE MODEL

- Proposer une méthode d'ensemble dressant un portrait en temps réel selon : texte, audio, vidéo
- Atteindre l'état de l'art dans chaque domaine

FIL ROUGE  
MOYENS  
NÉCESSAIRES



# REPARTITION ET COMPÉTENCES



Texte

Audio

Video

## Compétences clés :

- Deep Learning
- Architecture / Mise en Production
- Psychologie
- Droit / RGPD

# RESSOURCES

- Un accès à des machines AWS pré-configurées Deep Learning : m5a12x.large
- Fourni par Pôle Emploi
- Soutien d'un développeur web Pôle Emploi

# LIVRABLE

- Une plateforme Tensorflow.JS implémentant toutes les fonctionnalités



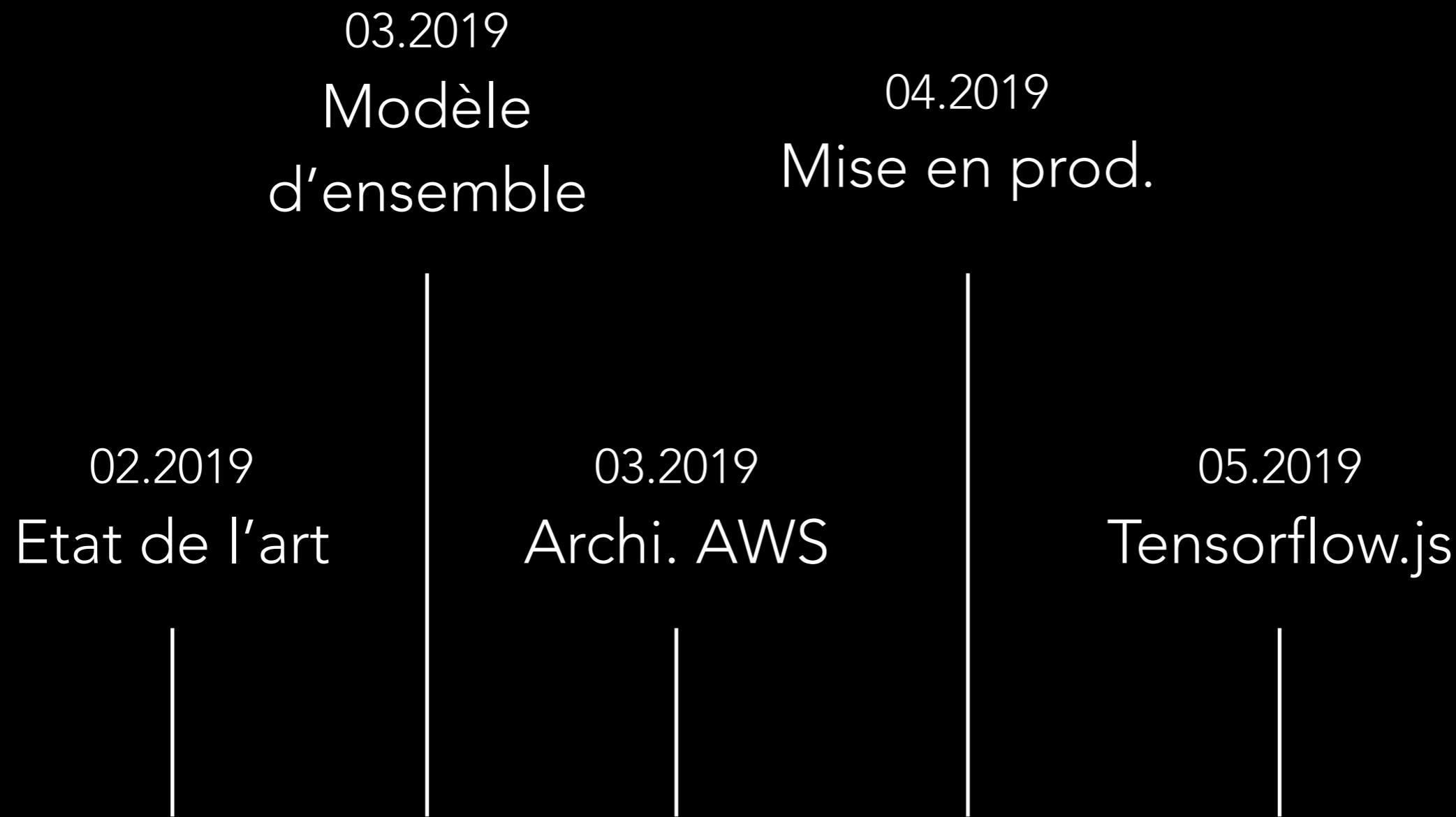
A WebGL accelerated, browser based JavaScript library for  
training and deploying ML models.

FIL ROUGE

PLANNING



# PLANNING



# DIFFICULTÉS

- Architecture AWS : Appui sur les développeurs Pôle Emploi
- Modèle d'ensemble : Appui sur l'encadrement universitaire

TELECOM PARISTECH - 2019

# MULTIMODAL EMOTION RECOGNITION

