# Artificial intelligence integration to crisis management. Part II

#### Matthieu François

GreenAl U.P.P.A. × Altanoveo

May 9, 2022





Matthieu François

#### 1 Crisis Management Context

- **2** Speech To Text Evaluation
- **3** Proposed System Architecture
- **4** Perspective of evolution

- \* ロ \* \* 個 \* \* 目 \* \* 目 \* - 目 \* - の < の

Matthieu François

GreenAl U.P.P.A. × Altanoveo

#### 1 Crisis Management Context

Overall & Technical Motivations Reminder Light structure

#### 2 Speech To Text Evaluation

B Proposed System Architecture

#### 4 Perspective of evolution

- \* ロ > \* @ > \* 注 > \* 注 > … 注 … のへの

Matthieu François

GreenAl U.P.P.A. × Altanoveo

# Crisis Management Context Overall & Technical Motivations Reminder

- Light structure
- 2 Speech To Text Evaluation
- B Proposed System Architecture
- 4 Perspective of evolution

- < ロ > < 個 > < 目 > < 目 > < 目 > < 回 > < の < の

Matthieu François

GreenAl U.P.P.A. × Altanoveo

#### **Overall Motivations**

An emergency requires **massive amount of data** to be processed with a **limited amount of time** and **limited human resources**.

- Use Articifial Intelligence (speech to text, NLP) for data indexing during a crisis
- Help crisis room in dealing with large amount of social network / broadcast media data
- Allow the crisis staff to understand the public reactions with respect to the crisis
- Using Altanoveo CWall screen as an information hub

Crisis Management Context	Speech To Text Evaluation	Perspecti
000000		

#### Technical Motivations

Evaluate our model

- 《口》 《聞》 《臣》 《臣》 三臣 - のへの

Matthieu François

GreenAl U.P.P.A. × Altanoveo

#### Technical Motivations

- Evaluate our model
- Deal with streaming flow
  - Flow acquisition
  - Inference time
  - Pipeline assembly

#### **Technical Motivations**

- Evaluate our model
- Deal with streaming flow
  - Flow acquisition
  - Inference time
  - Pipeline assembly
- Deal with Speech To Text output
  - No punctuation
  - Spelling mistakes

 Crisis Management Context Overall & Technical Motivations Reminder Light structure

2 Speech To Text Evaluation

B Proposed System Architecture

Perspective of evolution

- < ロ > < 個 > < 目 > < 目 > \_ 目 - うへの

Matthieu François

GreenAI U.P.P.A. × Altanoveo

Crisis Management Context	Speech To Text Evaluation	
Reminder		

• Project structure



Matthieu François

GreenAI U.P.P.A. × Altanoveo

Crisis Management Context		
0000000		

#### Reminder

- Project structure
- Few examples without any evaluation

Matthieu François

GreenAl U.P.P.A. × Altanoveo

Crisis Management Context			
0000000	000000	000000000000000000000000000000000000000	0000

#### Reminder

- Project structure
- Few examples without any evaluation
- Metrics & classification comparison

# Crisis Management Context Overall & Technical Motivations Reminder Light structure

- 2 Speech To Text Evaluation
- B Proposed System Architecture
- 4 Perspective of evolution

- < ロ > < 個 > < 目 > < 目 > < 目 > < 回 > < の < の

Matthieu François

GreenAl U.P.P.A. × Altanoveo

CwallAl system Surveillance de l'opinion par l'exploitation des médias & réseaux sociaux

GreenA



500

### Crisis Management Context

- Speech To Text Evaluation REPERE dataset Alpowermeter
- B Proposed System Architecture
- 4 Perspective of evolution

- イロ・ イロ・ イヨ・ イロ・

Matthieu François

GreenAI U.P.P.A. × Altanoveo



## Speech To Text Evaluation REPERE dataset

Alpowermeter

B Proposed System Architecture

4 Perspective of evolution

しょりゃくゆ そうやく ゆくえり

Matthieu François

GreenAl U.P.P.A. × Altanoveo

Crisis Management Context	Speech To Text Evaluation	
	00000	

Building on decades of research on broadcast news indexing

• [Giraudel et al., 2012], The repere corpus : a multimodal corpus for person recognition. LREC. 2012



GreenAl U.P.P.A. × Altanoveo

Crisis Management Context	Speech To Text Evaluation	
	00000	

Building on decades of research on broadcast news indexing

- [Giraudel et al., 2012], The repere corpus : a multimodal corpus for person recognition. LREC. 2012
- LREC: International Conference on Language Resources and Evaluation

Crisis Management Context	Speech To Text Evaluation	
	00000	

Building on decades of research on broadcast news indexing

- [Giraudel et al., 2012], The repere corpus : a multimodal corpus for person recognition. LREC. 2012
- LREC: International Conference on Language Resources and Evaluation
- 60 hours of annotated videos with audio transcriptions, overlaid text and people identification

Crisis Management Context	Speech To Text Evaluation	
	00000	

Building on decades of research on broadcast news indexing

- [Giraudel et al., 2012], The repere corpus : a multimodal corpus for person recognition. LREC. 2012
- LREC: International Conference on Language Resources and Evaluation
- 60 hours of annotated videos with audio transcriptions, overlaid text and people identification
- BFMTv and LCP channels : Broadcast news, debates, reports, People magazines

Crisis Management Context	Speech To Text Evaluation	
	00000	

Building on decades of research on broadcast news indexing

- [Giraudel et al., 2012], The repere corpus : a multimodal corpus for person recognition. LREC. 2012
- LREC: International Conference on Language Resources and Evaluation
- 60 hours of annotated videos with audio transcriptions, overlaid text and people identification
- BFMTv and LCP channels : Broadcast news, debates, reports, People magazines



#### ・ 《母 》 《臣 》 《臣 》 《 臣 》 《 臣 》

Matthieu François

GreenAI U.P.P.A. × Altanoveo

#### Speech To Text Evaluation

#### Word Error Rate Metric

$$WER = \frac{S+D+I}{N} \tag{1}$$

#### Speech To Text Evaluation

#### Word Error Rate Metric

$$WER = \frac{S+D+I}{N} \tag{1}$$

#### With

- S=number of substitutions
- D=number of deletions
- I=number of insertions
- N=number of words

GreenAl U.P.P.A. × Altanoveo

#### Speech To Text Evaluation

#### Word Error Rate Metric

$$WER = \frac{S+D+I}{N} \tag{1}$$

#### With

- S=number of substitutions
- D=number of deletions
- I=number of insertions
- N=number of words

 $WER(I \text{ love white dogs}, Why a love dogs) = \frac{3}{4}$  (2)

Matthieu François

GreenAI U.P.P.A. × Altanoveo



- Speech To Text Evaluation REPERE dataset Alpowermeter
- 8 Proposed System Architecture
- Perspective of evolution

- <ロ> <個> <ヨ> <ヨ> (ロ>)

Matthieu François

GreenAl U.P.P.A. × Altanoveo

Crisis Management Context	Speech To Text Evaluation	
	00000	

#### Alpowermeter

- Energy consumption tracking
  - on GPU in Watt
  - on CPU in Watt

-rançois

GreenAl U.P.P.A. × Altanoveo

Crisis Management Context	Speech To Text Evaluation	
	00000	

#### Alpowermeter

- Energy consumption tracking
  - on GPU in Watt
  - on CPU in Watt
- latency in s

Matthieu François

Crisis Management Context	Speech To Text Evaluation	
	000000	

#### Alpowermeter

- Energy consumption tracking
  - on GPU in Watt
  - on CPU in Watt
- latency in s
- Anticipate future hardware needs

#### Crisis Management Context

#### 2 Speech To Text Evaluation

#### 3 Proposed System Architecture

Speech Recognition Punctuation Named Entity Recognition Pipeline Consumption Server Side Key Entities Selection

#### Perspective of evolution

#### Matthieu François

GreenAl U.P.P.A. × Altanove

#### Crisis Management Context

2 Speech To Text Evaluation

#### 3 Proposed System Architecture Speech Recognition

Punctuation Named Entity Recognition Pipeline Consumption Server Side Key Entities Selection

#### Perspective of evolution

#### Matthieu François

#### Speech Recognition

# First: Channel stream audio arrives as a time series from a BFMTv / TF1 / LCP... flow.



ffmpeg audio acquisition

#### - 《日》 《聞》 《聞》 《理》 三目 - つんの

Matthieu François

GreenAI U.P.P.A. × Altanoveo

#### Speech Recognition with SpeechBrain

A wav2vec based model, SpeechBrain:

• We define a window inference size, to apply our model on.



- ▲日 > ▲雪 > ▲雪 > ▲ 目 > ④ ● ○

Matthieu François

GreenAl U.P.P.A. × Altanoveo

Crisis Management Context Speech To Text Evaluation 000000 Proposed System Architecture Perspec

#### Speech Recognition with SpeechBrain



- <ロ> <個> <ヨ> <ヨ> ヨ つへぐ

Matthieu François

GreenAl U.P.P.A. × Altanoveo

#### Speech Recognition with SpeechBrain

#### We concatenate our transcription according to the size of window.



l'ancien président américain obama a accusé samedi 23 octobre les républicains de menacer la démocratie avant une élection locale particulièrement serrée



Matthieu François

GreenAl U.P.P.A. × Altanoveo

#### Speechbrain Window Evaluation

When we vary the window size from 1 to 30 we can see that WER distribution won't change beyond 15s.





#### Speech Recognition with Vosk

We compared SpeechBrain to the Vosk Speach To Text model. It allows us to deal easier with streaming data.



l'ancien président américain obarna a accusé samedi 23 octobre les républicains de menacer la démocratie avant une élection locale particulièrement serrée

- \*ロ・\*園・\*画・\*画・ 一面・シック

Matthieu François

GreenAl U.P.P.A. × Altanoveo




l'ancien président américain obama a accusé samedi 23 octobre les républicains de menacer la démocratie avant une élection locale particulièrement serrée



Matthieu François

GreenAI U.P.P.A. x Altanoveo

Crisis Management Context Speech To Text Evaluation Proposed System Architecture Perspect

#### Speech Recognition Comparison

Median Word Error Rate distribution on REPERE per model per tv-show.



◆ロ → ◆母 → ◆臣 → ◆臣 → ● ● ● ● ● ●

#### Matthieu François

GreenAl U.P.P.A. x Altanoveo

Chais Management Context Speech to Text	Evaluation Proposed System Architecture	Perspecti
	000000000000000000000000000000000000000	

### Limits:

• Evaluation dataset is already cut, makes it easier to manage for SpeechBrain.



GreenAl U.P.P.A. × Altanovec

## Limits:

- Evaluation dataset is already cut, makes it easier to manage for SpeechBrain.
- Speechbrain: Window can cut the time serie in the middle of a word

## Limits:

- Evaluation dataset is already cut, makes it easier to manage for SpeechBrain.
- Speechbrain: Window can cut the time serie in the middle of a word
- Vosk: Easy to use API, hidden parameters

### Limits:

- Evaluation dataset is already cut, makes it easier to manage for SpeechBrain.
- Speechbrain: Window can cut the time serie in the middle of a word
- Vosk: Easy to use API, hidden parameters
- Vosk: 1.5Go size (38x lighter exists, but 2x worst on CV dataset)

26 / 58

#### In Practice

attaque spectaculaire dans les quartiers nord de marseille deux trains attaqués par une vingtaine de jeunes non ce n est pas westerne c est la réalité la mère des quartiers nord de marseille sera mon invité ce soir et puis le la guerre le conflit guérini mondour n est pas terminé guérini porte plainte contre mondour l avocat de jean noël guérini le président du conseil général des bouches durônes maître maisonneuvre et mon invité ce soir juste après les infos

#### Crisis Management Context

2 Speech To Text Evaluation

#### 3 Proposed System Architecture

Speech Recognition

#### Punctuation

Named Entity Recognition Pipeline Consumption Server Side Key Entities Selection

#### Perspective of evolution

#### - <ロ> <個> <ヨ> <ヨ> (四> <ロ>

Matthieu François

#### Punctuation

# Restore punctuation with a BERT base biLSTM, trained on french data.



Creen ALLI P. P. A. X. Altanoveo

Matthieu François

#### Crisis Management Context

2 Speech To Text Evaluation

#### 3 Proposed System Architecture

Speech Recognition Punctuation Named Entity Recognition Pipeline Consumption Server Side Key Entities Selection

#### Perspective of evolution

#### Named Entity Recognition

Apply Named Entity Recognition with a camemBERT based model (allowed us to deal with prediction score) or Spacy (allowed us to deal with POS).



GreenAl U.P.P.A. x Altanoveo



# We're still working on the Optical Character Recognition method. Especially to deal with advertising latency.



GreenAl U.P.P.A. × Altanovec

**Ukraine :** Une visioconférence des dirigeants du G7 se tient dimanche avec la participation de Volodymyr Zelensky

20.50 "DANS L'ENFER D'AZOVSTAL" L'ENQUÊTE

*"Ukraine Une visioconférence des dirigeants du G7 se tient dimancheavec la participation de Volodymyr Zelensky"* 

Matthieu François

GreenAl U.P.P.A. × Altanoveo



"À lire aussi sur bfmtvcom NON STOPÛ Que changerait une déclaration officielle de guerre par VPoutine"

Matthieu François

GreenAl U.P.P.A. × Altanoveo



- 《口》 《御》 《臣》 《臣》 三臣 - のへで

Matthieu François

GreenAI U.P.P.A. x Altanoveo



#### NER Efficiency

Distribution of number of Entities well recognized per file per Tv-programme.



▲口▶ ▲御▶ ▲臣▶ ▲臣▶ ―臣 \_ わえで

Matthieu François

GreenAl U.P.P.A. × Altanovec

#### In Practice

attaque spectaculaire dans les quartiers nord de marseille, deux trains attaqués par une vingtaine de jeunes. non ce n est pas westerne, c est la réalité, la mère des quartiers nord de marseille sera mon invité ce soir et puis le la guerre le conflit guérini mondour n est pas terminé, guérini porte plainte contre mondour l avocat de jean noël guérini, le président du conseil général des bouches durônes, maître maisonneuvre et mon invité ce soir juste après les infos

#### Crisis Management Context

#### 2 Speech To Text Evaluation

#### 3 Proposed System Architecture

Speech Recognition Punctuation Named Entity Recognition **Pipeline Consumption** Server Side Key Entities Selection

#### Perspective of evolution

GreenALLEPPA x Altanoveo

Matthieu François

#### Pipeline Consumption during 5 minutes inferences



# Vosk process requires almost 50% more CPU RAM than SpeechBrain.

Matthieu François

Artificial intelligence integration to crisis management. Part II

GreenAl U.P.P.A. × Altanoveo

39 / 58

#### Pipeline Consumption during 5 minutes inferences



#### CPU consumption distribution is quite similar between our 2 ASR

 Matthieu François
 GreenAl U.P.P.A. x Altanoveo

 Artificial intelligence integration to crisis management. Part II
 40 / 58



#### ASR represent 90% time of the pipeline, on a 10s window.

#### Pipeline Latency



# ASR represent 95% time of the pipeline and 140% of a 10s window size.

#### Crisis Management Context

#### 2 Speech To Text Evaluation

#### 3 Proposed System Architecture

Speech Recognition Punctuation Named Entity Recognition Pipeline Consumption

# Server Side

Key Entities Selection

### Perspective of evolution

#### · 《 曰 》 《 問 》 《 팀 》 《 팀 》 · [目 · 》) Q (?

Matthieu François

Crisis Management Context Speech To Text Evaluation		Proposed System Architecture	
			_

#### How to bring data to users?

- Run everything on server sequentially?
- Run everything on server with multiprocessing / threading?
- Run everything outside the server, with multiprocessing and feed a database?

Crisis Management Context	Proposed System Architecture	
	000000000000000000000000000000000000000	

#### How to bring data to users?

- Run everything on server sequentially?
- Run everything on server with multiprocessing / threading?
- Run everything outside the server, with multiprocessing and feed a database?

• Use WebSocket to discuss with server! Computer communication protocol that allowed us to send data between services over a single TCP connection.

Crisis Management Context	Proposed System Architecture	
	000000000000000000000000000000000000000	



#### -▲ㅁ▶ ▲卽▶ ▲끝▶ ▲끝▶ \_ 곁 \_ ∽) 옷(~)

#### Matthieu François

GreenAl U.P.P.A. × Altanoved

Crisis Management Context	Proposed System Architecture	
	000000000000000000000000000000000000000	



Artificial intelligence integration to crisis management	Part II	46 / 58

#### Crisis Management Context

2 Speech To Text Evaluation

#### 3 Proposed System Architecture

Speech Recognition Punctuation Named Entity Recognition Pipeline Consumption Server Side Key Entities Selection

#### Perspective of evolution

- 《日》《聞》《臣》《臣》 臣 うらの

Matthieu François

GreenAI U.P.P.A. x Altanoveo

Crisis Management Context Speech To Text Evaluation Proposed System Architecture Perspecti

#### Key Entities Selection

We can find: "emmanuel macron", "emanuell macron", "emmanuelmacron", "emmanuel makron" How can we consider them into only ONE entity?

GreenAI U.P.P.A. × Altanoveo

#### Key Entities Selection

### Levenshtein distance // WER metric



Matthieu François

GreenAI U.P.P.A. × Altanoveo

#### Key Entities Selection

## Levenshtein distance // WER metric



$$d < \min(n_i, n_j)/5 \tag{3}$$

With i and j two words, d the Levenshtein distance between them, and n the size of the word.

Crisis Management Context	Proposed System Architecture	
	000000000000000000000000000000000000000	



Crisis Management Context	Proposed System Architecture	
	000000000000000000000000000000000000000	



We can't count word occurrences forever.



Matthieu François

GreenAl U.P.P.A. × Altanoveo

#### Apply filter

We can't count word occurrences forever.

- We filter by time and count only words said in the last minute.
  - 1 minutes for context word
  - 5 minutes for entities

### Apply filter

We can't count word occurrences forever.

- We filter by time and count only words said in the last minute.
  - 1 minutes for context word
  - 5 minutes for entities
- We also filter on occurrences number to avoid too rare words.

Crisis Management Context	Proposed System Architecture	
	000000000000000000000000000000000000000	


### Enrich data

We added a twitter based data augmentation.

- Show the tweet number of the entity since we started the app.
- Verify the word spelling using last channel's tweet.

53 / 58

Crisis Management Context	Proposed System Architecture	
	000000000000000000000000000000000000000	0000

#### Server Side



latthieu François

GreenAl U.P.P.A. × Altanoveo

# Crisis Management Context

- 2 Speech To Text Evaluation
- **B** Proposed System Architecture
- **4** Perspective of evolution

#### - \* ロ > \* 御 > \* 画 > \* 画 > - 画 - のへの

Matthieu François

GreenAl U.P.P.A. x Altanoveo

Chisis Management Context Speech to Text L	valuation Proposed System Architecture	Perspecti
		0000 0000

#### Perspective

### ASR System improvements:

- Light vosk model
- Train / fine-tunning ASR model with REPERE
- Work more about hidden vosk parameters

Crisis Management Context		Perspecti
		0000

#### Perspective

### ASR System improvements:

- Light vosk model
- Train / fine-tunning ASR model with REPERE
- Work more about hidden vosk parameters

## **CWALL** application:

- Online representation of crisis
- Adding users practices reading from the CWALL

Thanks!



Matthieu François

GreenAl U.P.P.A. × Altanoveo

Crisis Management Context		Perspecti
		0000

#### References I

# [Giraudel et al., 2012] Giraudel, A., Carré, M., Mapelli, V., Kahn, J., Galibert, O., and Quintard, L. (2012). The repere corpus : a multimodal corpus for person recognition. LREC.

Matthieu François

GreenAl U.P.P.A. × Altanoveo