

Speech and Language Technologies for Security Applications

SALSA

Bernard PROUTS
salsa@vocapia.com

Février, 2015



SALSA

Objectif

Faciliter l'exploitation des données audio
dans un cadre judiciaire ou
de lutte contre le terrorisme et la criminalité

SALSA

Objectif

Faciliter l'exploitation des données audio
dans un cadre judiciaire ou
de lutte contre le terrorisme et la criminalité

ANR

Défi: "Liberté et sécurité de l'Europe, de ses citoyens et de ses
résidents"

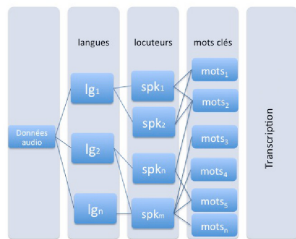
SALSA



volumes de données audio
en augmentation continue



Comment faire face
à toutes ces données ?



améliorer l'efficacité et réduire
la charge de travail des analystes

Enjeux

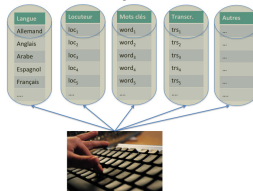
Données significatives pour l'enquête ?



Situation actuelle

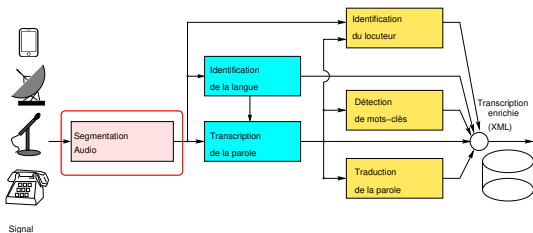


Objectif



Filterer les données en fonction de certains critères : langues, locuteurs, mots-clés...

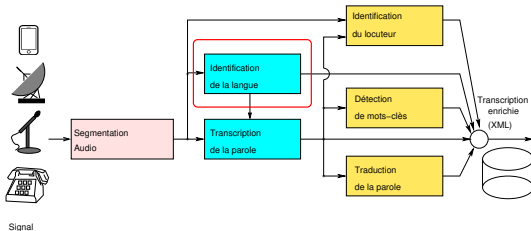
Technologies de traitement de la parole



segmentation audio : détecter la présence de parole dans un signal audio
 → filtrer les silences, les bruits... n'écouter que la parole (par exemple: HF)

```
<Channel num="1" sigdur="60.00" spdur="35.97"/>
...
<SpeechSegment ch="1" sconf="1.00" stime="1.16" etime="3.31" spkid="1"/>
<SpeechSegment ch="1" sconf="1.00" stime="8.65" etime="10.43" spkid="1"/>
...
```

Technologies de traitement de la parole



identification de la langue : quelle langue parle-t-on ?

détecter les changements de langue

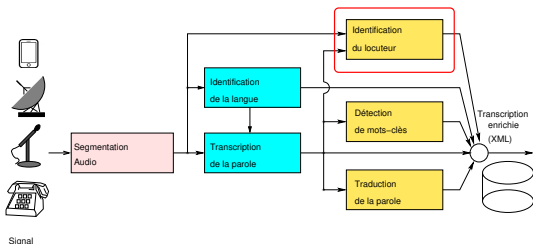
→ aiguiller l'appel vers le bon transcripneur (humain/machine)

→ aide à la caractérisation du locuteur

```
<Channel num="1" sigdur="60.00" spdur="17.05"/>
</ChannelList>
<SpeakerList>
<Speaker ch="1" dur="17.05" gender="2" spkid="1" lang="ita" lconf="0.88"/>
</SpeakerList>
<SegmentList>
<SpeechSegment ch="1" sconfs="1.00" stime="28.06" etime="28.99" spkid="1" lang="ita" lconf="0.88"/>
...

```

Technologies de traitement de la parole



segmentation et regroupement en locuteurs : qui parle quand ?

identification du locuteur : consiste à identifier un locuteur parmi un ensemble fini

→ identifier la voix d'un locuteur indépendamment du numéro d'appel utilisé

...

```
<SpeakerList>
```

```
<Speaker ch="1" dur="283.44" gender="1" spkid="MS2" name="Jean_Michel" conf="1.000"/>
```

```
<Speaker ch="1" dur="235.68" gender="1" spkid="MS3" name="Dupont_Theodor" conf="0.801"/>
```

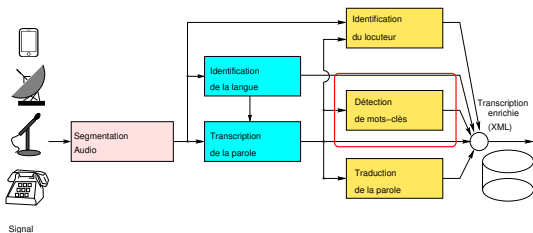
```
<Speaker ch="1" dur="176.65" gender="1" spkid="MS1"/>
```

```
<Speaker ch="1" dur="54.81" gender="1" spkid="MS4"/>
```

```
</SpeakerList>
```

...

Technologies de traitement de la parole



détection de mots-clés

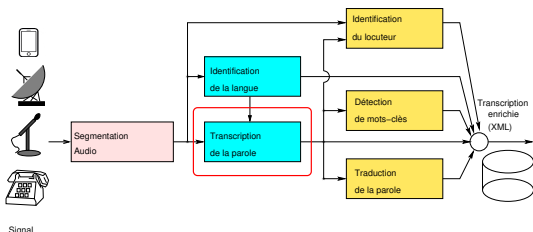
→ quels sont les documents contenant au moins un mot d'une liste définie (cannabis, came, drogue...)?

...

```
<SpeechSegment ch="1" sconf="1.00" stime="29.36" etime="39.52" spkid="1" lang="fre" lconf="1.00">
<Keyword stime="33.42" dur="0.16" conf="0.000053" id="KW107-4901"> cannabis </Keyword>
<Keyword stime="34.87" dur="0.15" conf="0.001092" id="KW107-7421"> drogue </Keyword>
</SpeechSegment>
```

...

Technologies de traitement de la parole



transcription de la parole

→ transcription automatique de la parole contenue dans l'audio

```

...
<SpeechSegment ch="1" sconfs="1.00" stime="4.34" etime="25.36" spkid="1" lang="fre" lconf="0.99" trs="1">
<Word stime="4.64" dur="0.34" conf="0.944"> allô </Word>
<Word stime="5.24" dur="0.10" conf="0.608"> -- </Word>
<Word stime="5.57" dur="0.15" conf="0.958"> oui </Word>
<Word stime="6.35" dur="0.21" conf="0.961"> oui </Word>
<Word stime="6.61" dur="0.33" conf="0.895"> Marina </Word>
<Word stime="7.17" dur="0.20" conf="0.961"> comment </Word>
<Word stime="7.37" dur="0.14" conf="0.963"> tu </Word>
<Word stime="7.51" dur="0.22" conf="0.959"> vas </Word>
...
    
```

Spécificités

- complémentarité des partenaires :
 - technologies de la parole
 - linguistique
 - investigations judiciaires
- lien étroit entre la recherche et les utilisateurs

Les résultats de SALSA faciliteront

- indexation d'importantes quantités de données
- détection des changements de langue ⇒ transcriptions
- extraction de toutes les conversations traitant d'un sujet donné

Innovations techniques

- **réduction du coût** de développement des systèmes
- adaptation des modèles pour cibler les **besoins spécifiques**
- développement de technologies gérant les **changements de langage**
- création d'une **interface** pour l'analyse de données audio

Innovations techniques

Vocapia Transcript Editor

Save

Help

Align

Translate

spec_scribe

fre

User: admin

Robert

bon ben ça réchauffe pas trop à la maison hein il fait froid encore dehors hein

Jacqueline

ah oui il fait froid -- qui a un temps je regarde là il y avait moins de cinq ce matin moins cinq à mon thermomètre hein -- et là il y a encore moins trois deux

Robert

alors je sais pas demain ça partait demain que ça remonte

Jacqueline

ah d'accord

Robert

ben parce que je l'ai baissé parce que j'ai dit comme ça va monter par le soleil alors parce que j'ai reçu tu sais j'ai reçu ma quittance d'électricité -- et j'ai regardé et tu sais c'est des prévisions deux fois sur trois

Jacqueline

oui

Robert

et et puis la troisième fois c'est c'est définitif -- alors j'ai regardé les prévisions sur deux moi et je j'ai consommé deux fois la prévision

Conclusions

La mise en oeuvre des technologies de la parole dans les applications sécurité permet de valoriser de nombreuses données audio actuellement sous-exploitées en leur donnant accès aux logiciels de traitement déjà disponibles pour le texte:

- indexation
- détection d'entités nommées
- détection de thèmes
- traduction

Technologies de la parole pour des applications de sécurité

Questions ?