

Modules technologiques de traitement de la parole

Bernard PROUTS, VOCAPIA Research

*Rencontres R&D Cap Digital
Paris 15 octobre 2013*

VOCAPIA
research

cap-digital
Paris Region



Identification de la langue et Transcription de la parole

- Description
- Intérêt
- Applications

- Segmentation audio, identification de la langue, transcription de la parole.
- Versions BS “broadcast speech” et CTS “telephone speech”
- Données d’entrée: signal audio AAC, AIFF, ASF, FLAC, MS-Wave, MPEG, Ogg/vorbis, Nist Sphere, Sun AU, ...)



- Plate-forme Linux x86 et x86_64 (Suse, Red Hat, OpenSuse, Fedora, CentOS, Ubuntu, Debian, ...)



- Logiciels proposés en service Web avec abonnement et facturation à la consommation ou au forfait
- Interfaces REST sur HTTPS
- Disponibilité: 24/7/365
- Redondance géographique (Roubaix, Strasbourg, Paris, Canada)
- Intégration simple (1h)
- Accès à une mise à jour quotidienne des modèles

Description (suite)

- 30 langues pour des données audio BS (radio/TV)
- 50 langues pour des données audio CTS (telephone)
- outils de création de modèles de nouvelles langues à disposition de l'utilisateur

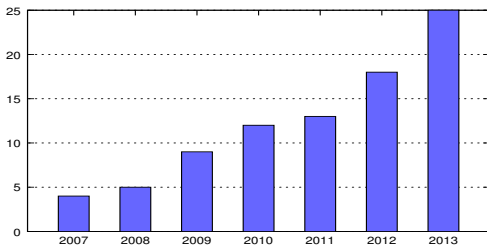
Intérêt

- Tri de données multilingues

Applications

- Défense
- Centres d'appel
- Diffuseurs radio/TV

25 langues BS



<2008	4	anglais, français, arabe, mandarin
2008	5	russe
2009	9	espagnol, allemand, italien, néerlandais
2010	12	grec, polonais, finnois
2011	13	portugais
2012	17	bulgare, roumain, tchèque, ukrainien
2013	25	turc, slovène, danois, hongrois, lituanien + 3 langues

7 langues CTS français, anglais, espagnol, italien, russe, mandarin, arabe

Intérêt

- 1 Valoriser les données audio en leur ouvrant l'accès aux logiciels de traitement disponibles pour le texte
 - Indexation
 - Détection d'entités nommées
 - Détection de thèmes
 - Résumé ...
- 2 Réduire les temps et les coûts de production de transcriptions exactes
 - Sous-titrage
 - Transcription de téléconférences
 - Transcription de séances parlementaires
 - Comptes-rendus de réunion

Applications

- Défense
- Pige des médias
- Speech analytics pour les centres d'appel
- Services de transcription

Réduction des coûts de sous-titrage (NPO TV)

- 20% des programmes sont trop difficiles pour l'état de l'art
- pour 60% des programmes le temps de traitement manuel est réduit de 50%
- pour 20% des programmes le temps de traitement manuel est réduit de 70%
- la qualité de la ponctuation et de la casse est primordiale
- le taux de confiance moyen permet de mesurer la difficulté du document

- Téléconf de sociétés du Fortune 500 (bilans trimestriels)
- Temps de latence réduit (traitement en quasi temps réel)
- Service SaaS 24/7
- Partenaire américain

