



SEEKIOTECH

FROM SPEECH TO TEXT

Speech Recognition Systems
for Automatic Transcription,
Voice Command & Dialog applications

Frédéric Beaugendre

www.seekiotech.com

SeekioTech

- Start-up hosted at the multimedia incubator « la Belle de Mai », in Marseille
- Relocate at CREATIVA, Avignon, end of 2013



BUSINESS ACTIVITY

- Design and Distribution of Speech Recognition Systems

COMPANY VALUES

- Tailored speech solutions to specific needs
- Dynamic adaptation to the task and environment

ADDED VALUE

- Speech as a new interface
- Robust systems, better adapted → better accuracy

VOICE COMMAND

Voice-Picking
Home Automation
Voice Command for Disabled
...



Command & Control
System



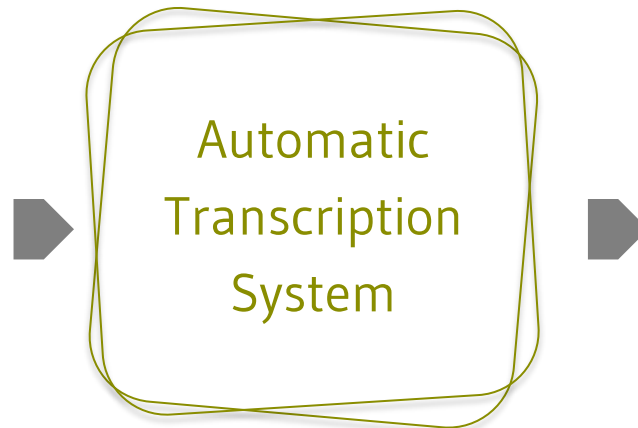
- "Yes"/"No"
- A pack of Coke
- Set the temperature to 20 degrees"

DIALOG SYSTEMS



TRANSCRIPTION SYSTEM

- Full text transcription
- Text alignment
- Indexation of audio/video context



<Turn speaker="S108M">
<Sync time="55.31"/>

An exceptional event in opening of this
news program...

TRANSCRIPTION SYSTEM

Application Domains & Target users



- Broadcasters (e.g. television and radio news), multimedia web content, digital archiving, economic intelligence (automatic indexation of companies' audio/video content)



- Automatic transcription of meetings, conferences, university supports



- Telephony: automatic transcription of telephony calls for call-centers in order to ease customer satisfaction analysis



TRANSCRIPTION SYSTEM

Functionalities/Hilights

- Above 95% accuracy on broadcast news content
- On-the-fly language models adaptation
- Segmentation of acoustic flow into coherent segments (music, microphone vs telephony speech, music + speech)
- Speaker diarization (segmentation, tracking, gender detection)
- Speaker identification (requires specific model training for the related speakers)
- Language detection
- Complete transcription of the speech parts into text
- Confidence scores at the word and sentence level
- Punctuation

INDEXATION OF MULTIMEDIA CONTENT (AUDIO/VIDEO)



rtbf.be Recherche Média Graphe Paramètres **transmedia player**

Sélectionnez un filtre global :

- Date
- Programme
- Texte

Sélectionnez un type de sujet :

- Personne
- Lieu
- Activité
- Objet
- Organisation
- Abstrait

Elio Di Rupo

Glissez pour supprimer

Belgique : la note d'Elio Di Rupo - explications de Thomas Gadisseux et de Jean Faniel
JT La Une 04/07/2011 03:03

Belgique : la note d'Elio Di Rupo - le volet institutionnel
Belgique : durant 1heure et demi, Elio Di Rupo a révélé à la

JT La Une 04/07/2011 01:47

La note d'Elio Di Rupo - itw de Charles Michel
Elio Di Rupo a présenté sa note à la presse.

JT La Une 04/07/2011 00:43

Belgique : la note d'Elio Di Rupo - explications de Thomas Gadisseux et de Jean Faniel
Belgique : durant 1heure et demi, Elio Di Rupo a révélé à la

JT La Une 04/07/2011 02:06

Laeken : arrivée d'Elio Di Rupo
Laeken : le formateur Elio Di Rupo est arrivé au Palais afin de

JT La Une 04/07/2011 00:16

Info du jour : la note d'Elio Di Rupo
Belgique : durant 1heure et demi, Elio Di Rupo a révélé à la

JT La Une 04/07/2011 00:16

Belgique : la note d'Elio Di Rupo - le volet socio-économique
Belgique : durant

JT La Une 04/07/2011 01:50

Réforme de l'état : La note du formateur Elio Di Rupo
Réforme de l'état : La

JT La Une 06/07/2011 01:52

TRANSCRIPTION ACCURACY

Combination of different parameters:

- Coverage of the lexicon
- Quality of the language models (good coverage of the word transition probabilities)

1 Missing word in
the lexicon



Several neighboring
mistakes (bi-
gram/tri-gram
language models)



DYNAMIC ADAPTATION

- « Named Entities » are extremely important for news/content indexation, search, classification
- General Language Models contain a finite number of words (most common)
- Typical size of a large vocabulary lexicon: hundreds of thousands of words, in a specific time frame and for a specific domain (or combination of domains)

DYNAMIC ADAPTATION

Language Models are dependent on Application Domains!!

- Depending on the application context, the named entities to be spotted will be completely different:
 - News: personalities, institution, journalists, etc..
 - Town council: name of the elected, city districts, etc...
 - Company meeting: company projects/products, the employees, ...
- How do we manage brand new named entities appearing in the news?
- How do we manage names entities related to old archived?

DYNAMIC ADAPTATION

SeekioTech Approach

- Very Large Vocabulary Language Models
- Mostly Based on Broadcast news content
- On-the-fly adaptation using textual content:
 - Wikipedia/Dbpedia on the topic
 - Specific Data Internal to a company
 - Text included in slides for a conference
 - News from the day (from news agencies, online newspapers or internal specific database)

A posteriori adaptation: OTMedia Project

- 1st Recognition Pass
- Extraction of linked articles in the OTMedia Database
- Dynamic adaptation of the language models
- 2nd Recognition Pass

The French Transmedia Observatory

L'observatoire Transmedia
Marie-Luce Viaud / mlviaud@ina.fr



<http://www.youtube.com/watch?v=N9GyfxbAJ2o&feature=plcp>



EXEMPLE 1 : RECENT NEWS

Boston Marathon bombings

Boston Marathon bombings

From Wikipedia, the free encyclopedia

Coordinates: 42°20′59.46″N 71°4′43.93″W﻿ / ﻿42.3497222°N 71.0788722°W﻿ / 42.3497222; -71.0788722

"Boston bomber" redirects here. For the Douglas Boston bomber aircraft, see Douglas A-20 Havoc.

During the Boston Marathon on April 15, 2013, two pressure cooker bombs exploded at 2:49 pm EDT (18:49 UTC), killing 3 people and injuring an estimated 264 others.^[3] The bombs exploded about 13 seconds and 210 yards (190 m) apart, near the finish line on *Boylston Street*.^{[2][8]} The Federal Bureau of Investigation (FBI) took over the investigation, and on April 18, released photographs and surveillance video of two suspects.^{[9][10][11]} The suspects were identified later that day as Dzhokhar and Tamerlan Tsarnaev.

Shortly after the FBI released the images, the suspects allegedly killed an MIT police officer, carjacked an SUV, and initiated an exchange of gunfire with the police in *Watertown, Massachusetts*. During the firefight, an MBTA police officer was injured but survived, in spite of severe blood loss. Tamerlan Tsarnaev was run over by his brother Dzhokhar, and Dzhokhar Tsarnaev was injured but escaped. An unprecedented manhunt ensued on April 19, with thousands of law enforcement officers searching a 20-block area of *Watertown*.^[12]

During the manhunt, the authorities asked residents of *Watertown* and surrounding areas, including *Boston*

to stay indoors, and the public transportation system and most businesses were shut down, resulting in a deserted urban environment of historic size and duration. After the "shelter-in-place" advisory was rescinded, a *Watertown* resident in a boat in his back yard. Dzhokhar was arrested and taken to a hospital ship.

During an initial interrogation in the hospital, Dzhokhar—who was arrested without a warrant—said Tamerlan was the mastermind. He said they were motivated by the wars in *Iraq* and *Afghanistan*, and that they were self-radicalized and members of terrorist groups, but that they had learned how to build explosive devices from a Qaeda affiliate in *Yemen*.^[16] He said that he and his brother had decided to travel to *New York City* to bomb *Times Square*. Dzhokhar was charged on April 19 with use of a weapon of mass destruction and malicious destruction of property. He has pleaded not guilty to 30 total charges.



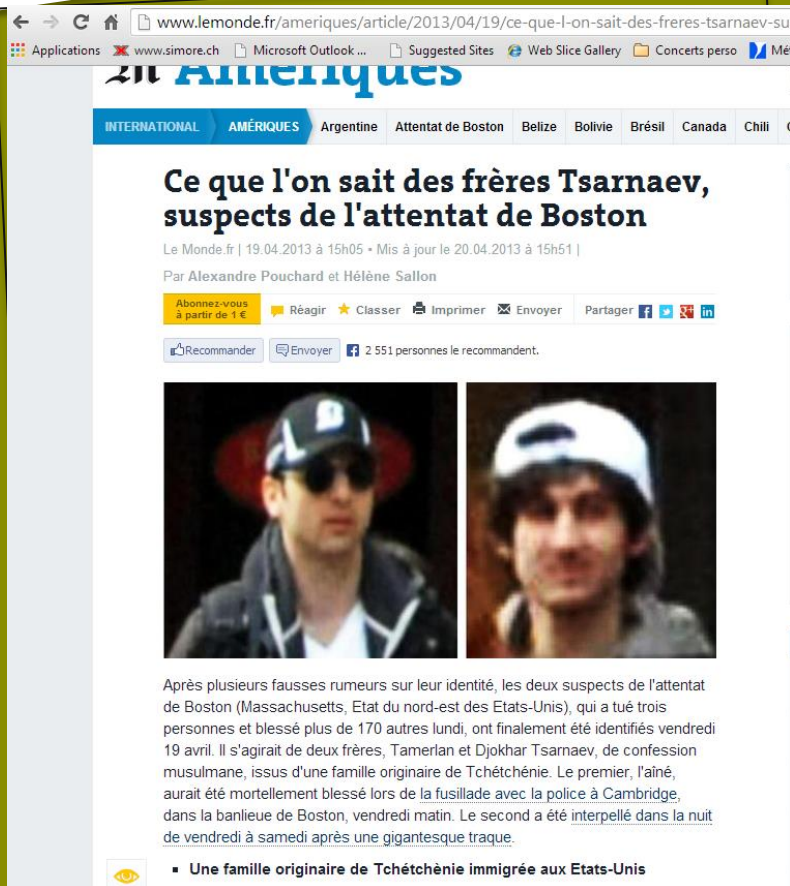
YouTube video player interface showing a video titled "Boston Marathon Bombings - Clear footage" by GlobalLeaks. The video has 735,189 views and 29,924 subscribers. The video content shows a large explosion at the Boston Marathon finish line, with a massive plume of white smoke rising from the street. The video player includes a progress bar, volume control, and a share button.

Facebook video player interface showing a video titled "Etats-Unis: Tamerlan Tsarnaev finalement enterré" (Official) by Euronews. The video has 8 views. The video content shows a large crowd of people gathered for a funeral procession, with many people holding flags. The video player includes a progress bar, volume control, and a share button.

DYNAMIC ADAPTATION

Use of textual data in order to allow on-the-fly language model adaptation:

- **News from the day**
 - from news agencies
 - online newspapers
 - internal specific database



The screenshot shows a web browser displaying a news article from Le Monde.fr. The article title is "Ce que l'on sait des frères Tsarnaev, suspects de l'attentat de Boston". The article is dated 19.04.2013 and is written by Alexandre Pouchard and Hélène Sallon. Below the title, there are two photographs of the suspects, Tamerlan and Dzhokhar Tsarnaev. The article text mentions that the suspects were identified on April 19th and were brothers from Chechnya. A bullet point at the bottom of the article reads: "▪ Une famille originaire de Tchétchénie immigrée aux Etats-Unis".



EXEMPLE 1 : REGIONAL VIDEO CONTENT



YouTube FR

Nous avons défini la langue que vous préférez

Voici votre langue préférée : Français. Vous pouvez modifier ce paramètre ci-dessous.

LCM
Une chaîne TV SUD

0:02 / 1:36

Le pavillon M, vitrine de MP2013

LCM - 10 736 vidéos 352 vues

 S'abonner 2 872

J'aime À propos de Partager Ajouter à

Publiée le 12 déc. 2012
Gros plan sur le pavillon M situé sur la place Bargemon, une structure architecturale éphémère qui a pour vocation de vitrine pour Marseille Provence 2013.

Plus

DYNAMIC ADAPTATION

Use of textual data in order to allow on-the-fly language model adaptation:

- Description of the video in a specific field
- Description in a database (e.g. dbpedia)
- Description in an event related web site



Publiée le 12 déc. 2012
Gros plan sur le pavillon M situé sur la place Bargemon, une structure architecturale éphémère qui a pour vocation de vitrine pour Marseille Provence 2013.



PARTENAIRES



Thank you for
you attention !!

Frédéric BEAUGENDRE

C.E.O

Pascal NOCERA

Head of Research & Development

INFO : info@seekiotech.com

SALES : sales@seekiotech.com

TEL : +33 (0) 6 95 83 69 83

WEB : www.seekiotech.com



www.seekiotech.com