Audio Hot Spotting

Qian Hu

781-271-2959 • qian@mitre.org

MITRE Sponsored Research

MITRE Technology Program

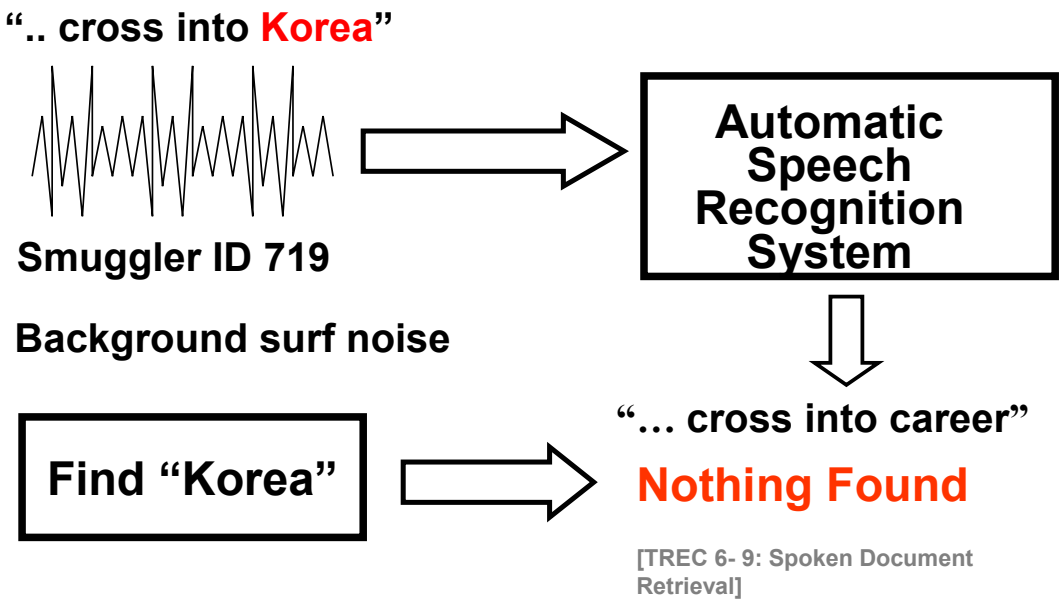
Problem

Nuggets of information are buried within huge mountains of multimedia data. A system is needed to quickly and automatically identify and retrieve these audio "hot spots."

The current approach to audio information retrieval of simply combining text-based information retrieval with automatic speech recognition does not meet user needs in real applications.

Background

Current Approach: ASR + IR Doesn't Meet the Needs of Critical Applications



Objective

- Provide prototype hot-spotting capability to enable efficient audio filtering and retrieval
  - Determine the limitations of algorithms in existing component technologies
  - Leverage, integrate, and extend the best technologies for the hot-spotting problem
  - Research and develop new algorithms when COTS and GOTS systems fail to meet the requirements of audio hot-spotting
  - Research and develop audio-specific query algorithms making use of multiple types of audio information

Activities

- **Speech Recognition**: Evaluated COTS/GOTS with real application data; experimented with language model adaptation
- **Speaker Identification**: Extended GOTS algorithms to improve "speaker-change" detection in domains including several unknown speakers
- Experimented with query expansion techniques to compensate for transcription errors
- Extended Audio Hot Spotting prototype to:
  - support foreign languages
  - search across multiple "audio documents"
  - search documents on reusable, predefined text queries

MITRE

Highlight

Input

- Real Media
- WAV Files
- AVI
- MPG
- MPA
- MPEG
- MP4
- MP3

Preprocessing

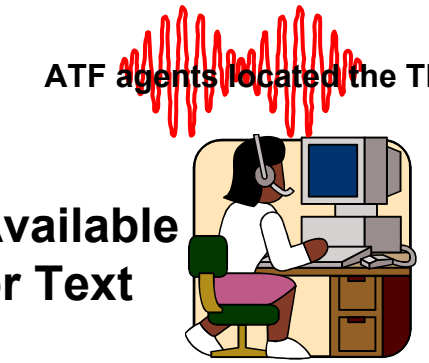
- Speech Recognition / *LM experimentation*
- Speaker ID / *Speaker Change*
- *Automatic Keyword Indexing*
- *Data Structuring*
- *Audio Feature Extraction*

Output

- Time Indexed Text
- Time Indexed Speakers & Speaker Statistics
- Lexical Index List
- Structured Data in Relational Database
- Time Indexed Audio Features

Blue = MITRE

Red = GOTS/COTS



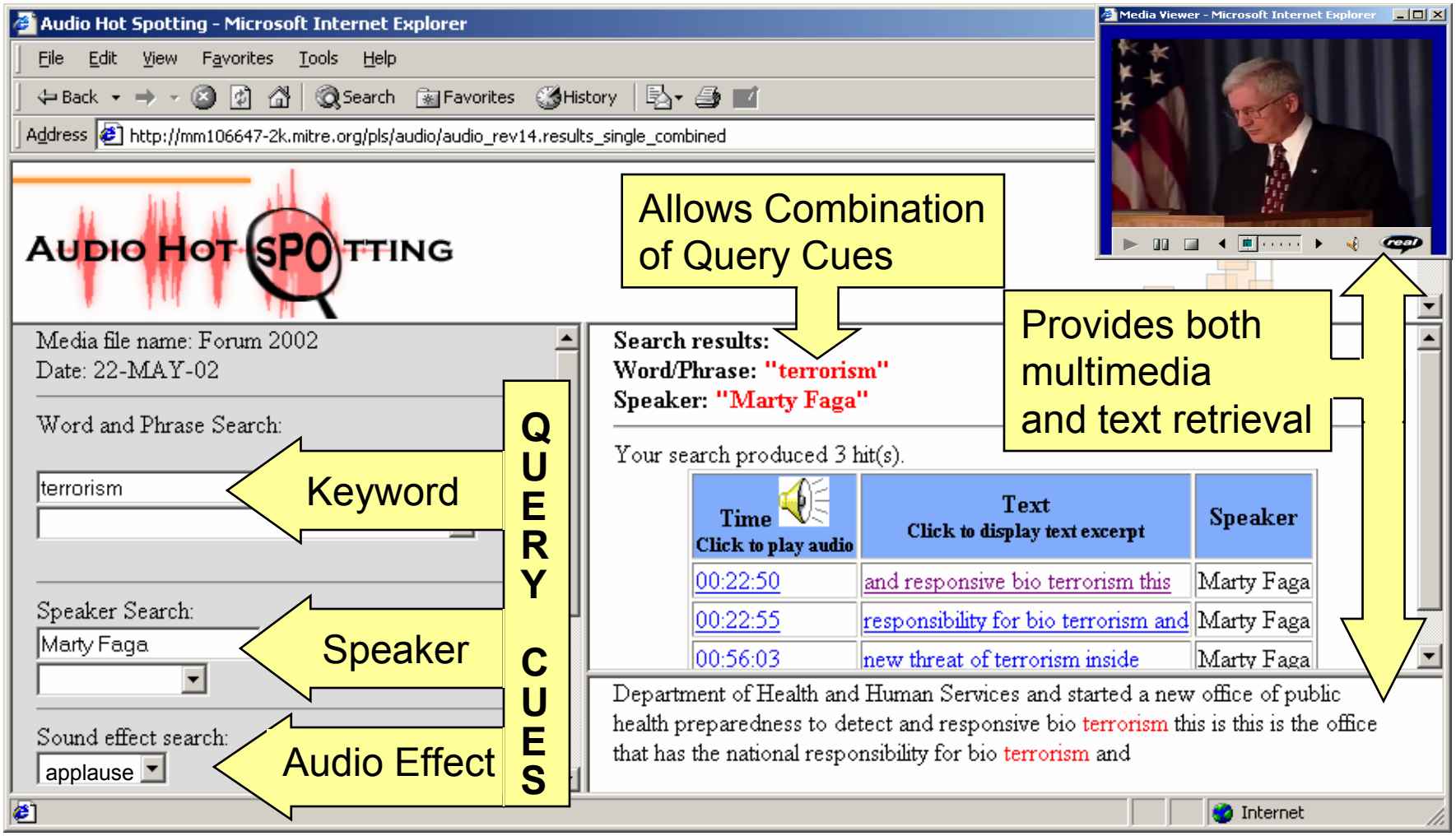
- *Query Language*
- *Hot Spotting GUI*
- *IR for High Error Rate Speech*

Identified Selection Available as Original Media or Text

MITRE

Demonstration

Built Audio Hot Spotting prototype to query and retrieve key words, speakers, and some audio effects for both English and Spanish



MITRE

Impacts

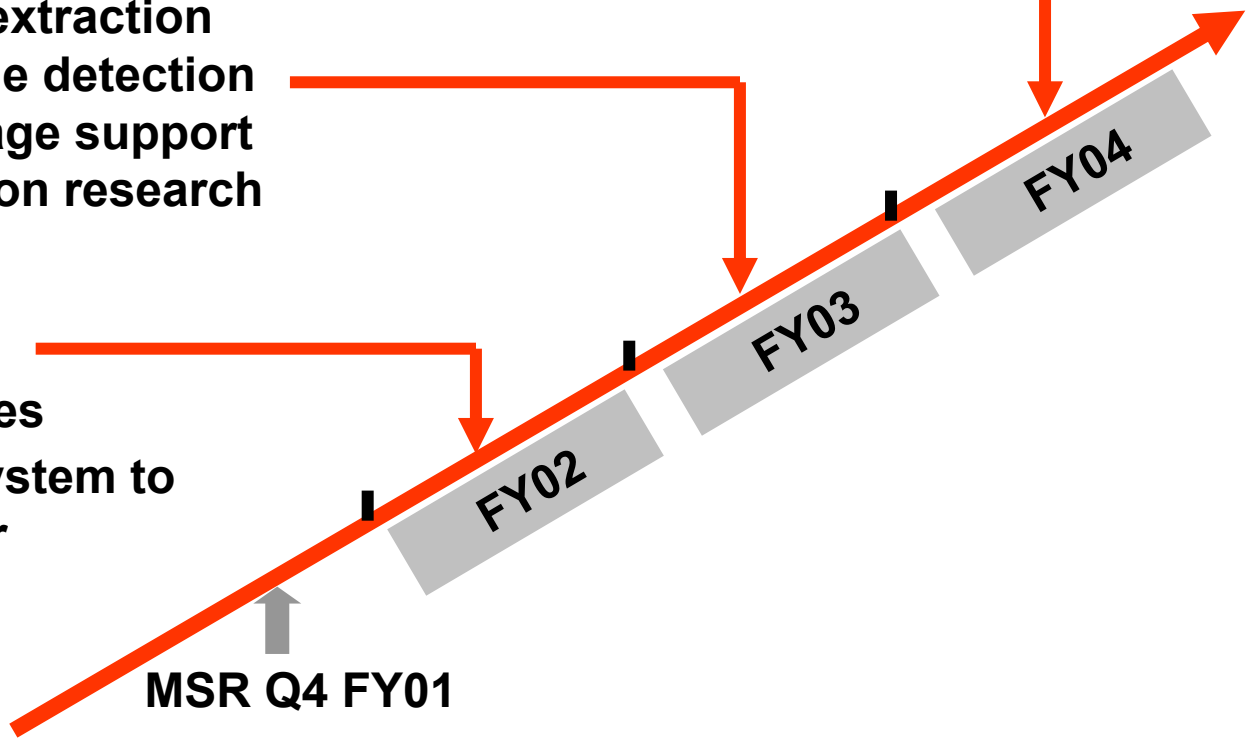
- Presented our work at the 5th International Conference on Text, Speech and Dialogue
- Presented a paper at 2002 Fall IEEE Conference on Technologies for Homeland Security
- Briefed government organizations, including NSA, DARPA, ARDA, FBIS, and AFRL
- Worked with industry leaders to improve component technologies

Future Plans

- Audio-specific query algorithms
- Domain-specific applications
- IR & audio mining research
- Explore multiple ASR engines

- Audio feature extraction
- Speaker change detection
- Foreign language support
- Query expansion research

- Integrate and extend component technologies
- Develop a prototype system to hotspot speaker and/or keywords



MITRE