



STS Media Xplorer

Empower your mobile operators

The need to enhance your human force with audio and voice processing technologies is critical in today's world. It is also necessary to provide them with tools to deal with the ever growing volume of audio/video data. What about producing actionable intelligence while it is still relevant? The problems are further multiplied when dealing with low quality audio collections such as radio or poor cellular traffic, multiple languages, and you are limited with size, weight, and power (SWAP) resources. Automated speech processing tools have made some advancements, but they are still very far from being able to replicate the human mind or ear (or even doing significant processing on an untethered machine). So how do you empower your operators and magnify the ability?

STS Media Xplorer enables defense agencies to tap into the massive amounts of audio and video communications, at unparalleled speeds, with the highly accurate search results and with the most flexible analytic capabilities available. The tool can be set up to automatically process large amounts of recorded audio anywhere from low quality cellular traffic and radio communications to wiretaps or even voicemail, making the content searchable within seconds. By using this unique phonetic approach, it eliminates the need to translate from speech to text, and then mine that text. All this power can be available on any commodity Windows desktop or laptop.

The tool is designed for flexibility. It does allow for further integration, but was mainly built for standalone operations. It offers full functionality without touching any external systems or networks.

Our focus here was to reduce any reliance on backend systems or analysts and increase the speed at which decisions could be made. There are two primary components to the approach: the ability to **classify audio** into manageable information buckets by language, group, or speaker AND the ability to rapidly and accurately **search** across one or many of those buckets. This all can be accomplished within one simple to use interface.

The unique classifier framework employed can identify and support any language, speaker, or group. This is done using the raw audio, not using Metadata provided from the sensor. If Metadata is available, however, it can be ingested and also used as part of the search criteria.

It is possible to search any spoken word(s), enabling the timely identification of relevant threats and empowering agencies to respond immediately. There is no dictionary to maintain, so searches on proper names, inexact spellings, slang, etc. can be done with extremely accurate results. Word spotting is really the most basic concept and there is so much more to this solution. It is even possible to produce a conversational "gist", using just the phonetic search approach that is outside of using a speech to text (STT) engine.

The underlying technology is language agnostic, so if the language required is not one of the 40+ languages currently supported, it could always be built. Constructing new language packs will take some time and money, however depending on the mission and requirements; it is possible to use the multilingual "universal" language pack that supports **ANY** and **ALL** spoken languages for search.

Searching phonemes, the smallest unit of human speech, is still relatively new. All other systems use an approach that many are familiar with where they try to convert speech to text and then mine the text. This "phonetic" approach is different; it is based on years of research as well as an understanding of the operators, environment, and mission. It does not provide transcript output (which tends to miss much of the key information), but it does provide an extremely rapid and accurate result based on the information you are looking for. It is far more robust to audio variances such as a speaker's language, accent, dialect, gender and age which enables timely identification of cover and threat terms. The tool allows users to efficiently sift through large amounts of data and find the information that is only relevant to them and the mission.

