FREQUENTLY ASKED QUESTIONS ON
AUDIO MINING

SUMMARY ///

Audio mining embraces speech recognition, keyword, phonetic or transcription technologies to extract insights from prerecorded voice streams. This insight can then be used to classify calls, trigger alerts/workflows, and drive operational and employee performance across the enterprise. This speech analytics solution is commonly used in contact centers to mine the customer interactions to provide valuable insights into products, services, and processes to help reduce operational costs and improve customer satisfaction.

Uniphore Software Systems, a pioneer in providing voice-based mobility solutions has integrated its audio mining solutions for many enterprises to improve data processing, customer satisfaction and fraud identification. This white paper outlines the frequently asked questions related to audio mining, its credibility, enrollment, sample use cases, and more to give a fair understanding on integrating audio mining in contact centers.
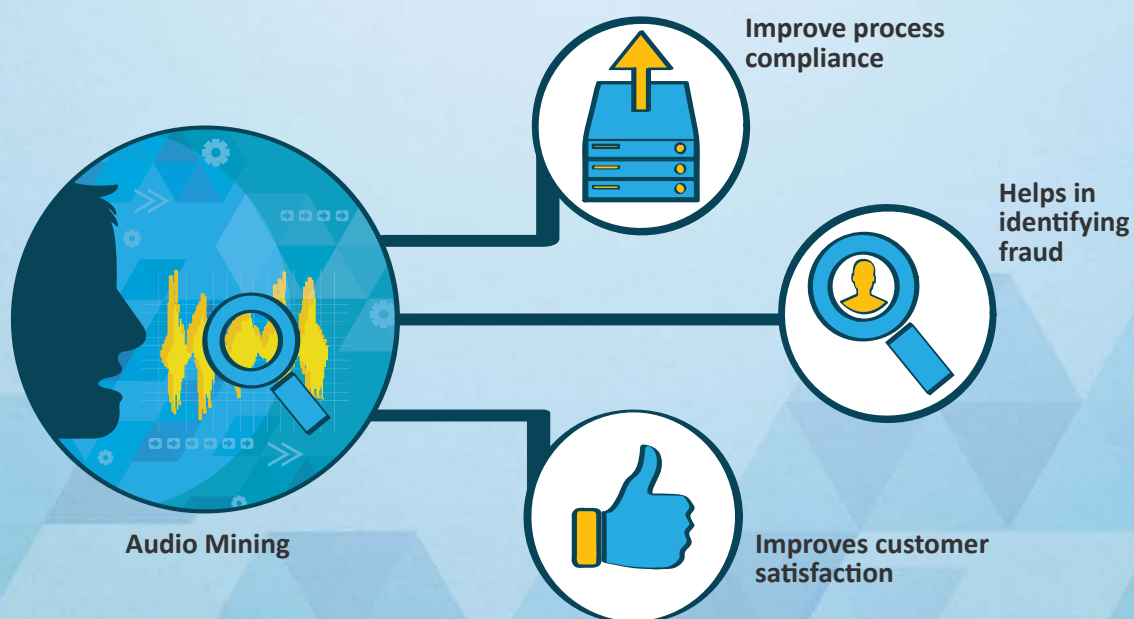
Improve process compliance

Helps in identifying fraud

Improves customer satisfaction

Audio Mining

uniphone
Speech-based Mobility Solutions

TABLE OF CONTENTS ///

uniphore
Speech-based Mobility Solutions
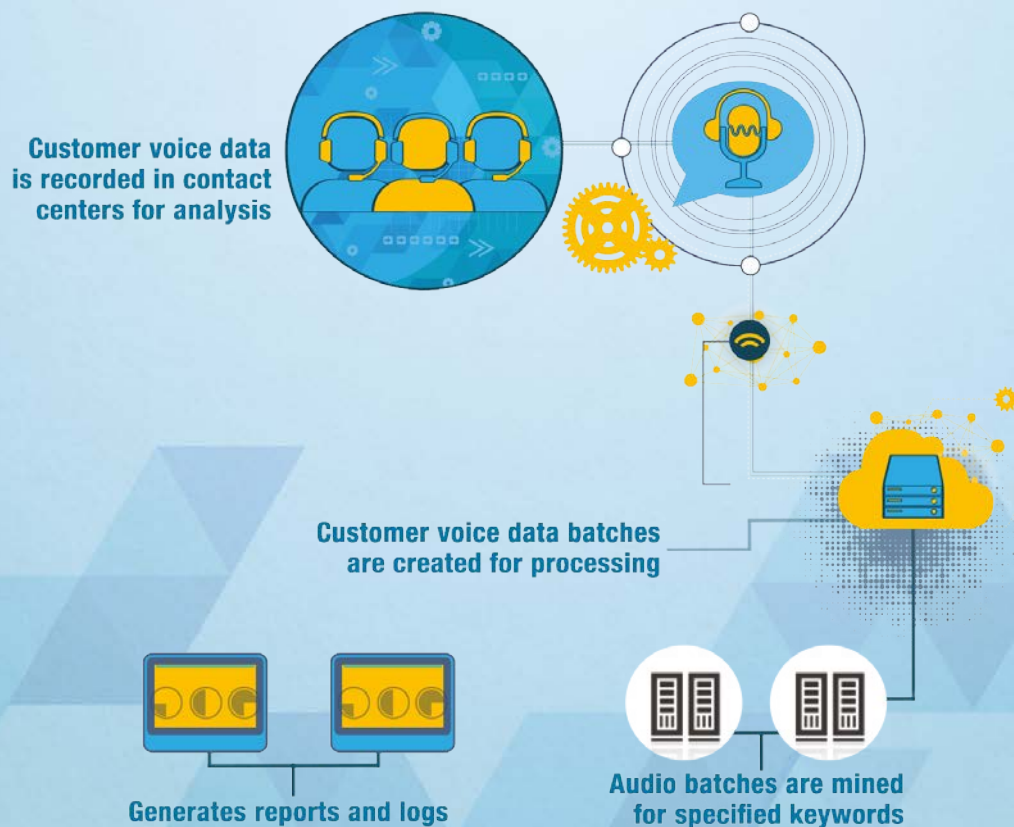
GET STARTED WITH AUDIO MINING ////

What is Audio Mining?

Speech recognition solutions for contact centers use techniques commonly referred to as audio mining, where large volumes of audio are searched for occurrences of specific words or phrases. Mining recorded customer interactions can provide valuable insight into products, services and processes to help reduce costs and improve customer satisfaction.

How does Audio Mining work?

The process begins with speech recognition engine ingesting a large number of recorded conversations. Using the audio mining and indexing technologies described above, the solution recognizes words within this large volume of unstructured information and organizes them into user-created and self-suggesting categories.

The software can accomplish this because it "understands" the content. For instance it might sort the recordings into three categories: *(1)* customer complaint calls, *(2)* calls in which a new product offering is mentioned, and *(3)* calls in which a competitor is named during the conversation. The solution drills deeper into each category and identifies clusters of calls with commonalities that suggest a root cause. It's this part of the solution that makes the biggest business difference. The key factor for success in terms of business impact lays in the quality and depth of the searchable index and the precision of the categorization and root-cause identification routines.

Customer voice data
is recorded in contact
centers for analysis

Customer voice data batches
are created for processing

Generates reports and logs

Audio batches are mined
for specified keywords

uniphore
Speech-based Mobility Solutions

Why enterprises need audio mining?

In contact centers large and small, the sheer volume of recorded interactions makes analyzing captured conversations too difficult. The inability to systematically analyze large amounts of audio data causes massive challenges:

POOR QUALITY CONTROL

Poor customer experience, and inability to analyze trends of representative misconduct

MISSED OPPORTUNITIES FOR EXTRA REVENUE

Inability to identify relevant Cross-Sell, Up-Sell opportunities on a real-time basis

HIGHER RISK OF FRAUD

Inability to identify mis-selling and improper conduct in real-time contact center interactions

How are enterprises trying to solve these problems?

To get control of tons of audio data, enterprises currently devote extensive funds towards suboptimal solutions:

Transcribe audio data

The limitations in this method are:

- *Long delays between time audio is recorded and time it is transcribed and analyzed*

- *Extensive time and effort required for transcription*

Manual re-listening

The limitations in this method are:

- *Often inaccurate and inconsistent (one reviewer thinks a customer is satisfied, another doesn't)*

- *High level of human resource efforts required*

Using voice loggers in IVR to analyse the audio data for finding the 'keywords'

- *Low accuracy rates*

- *Language dependant*

UNIPHORE'S AUDIO MINING SOLUTION

Why Uniphore's audio mining solution?

Uniphore's audio mining solution analyzes human speech to extract useful information about the content. The application identifies particular keywords or phrases, extracts them, and analyses them for critical business insights.
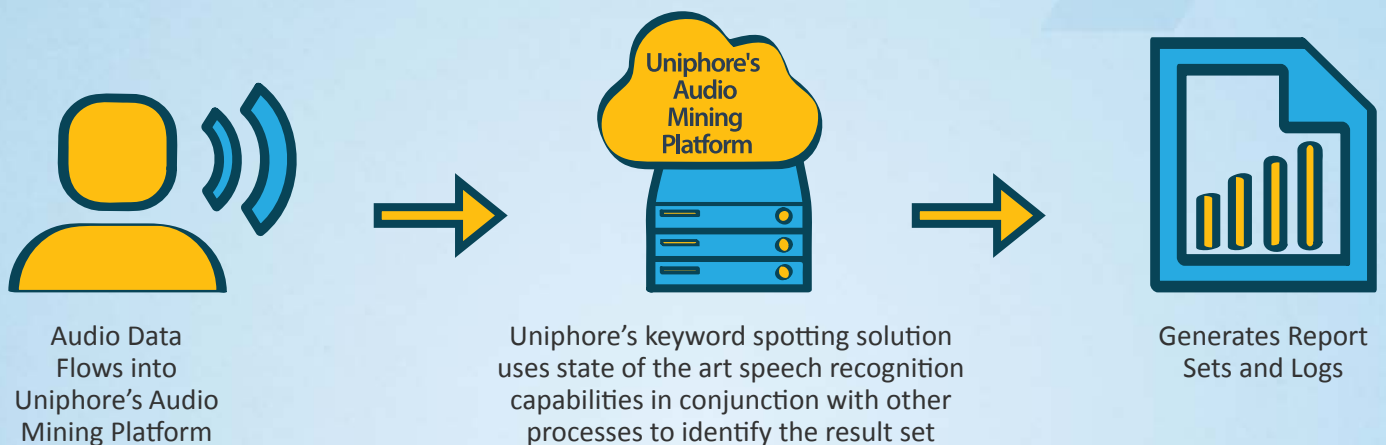
Uniphore's audio mining solution works in simple 3 steps:

Audio Data Flows into Uniphore's Audio Mining Platform

Uniphore's keyword spotting solution uses state of the art speech recognition capabilities in conjunction with other processes to identify the result set

Generates Report Sets and Logs

What are the features of Uniphore's audio mining solution?

The key features of Uniphore's audio mining solution are:

- *Text-independent*

- *Language and accent-independent*

- *Unsurpassed state-of-the-art accuracy – consistently achieving above 95%.*

- *Supports single/multi-speaker calls (2/4 wire)*

- *Web service interfaces for quick, easy, and secure integration*

- *Integrated security and role-based authorization*

- *Multi-tenancy, scalability, and multi-site high availability*

- *Web-based management applications for security, system administration, and domain experts.*

- *Advanced Noise Reduction and Silence Deduction*

- *Operational in Real-time or Batch-mode*

- *On-Premise or Cloud based Delivery model*

- *S-a-a-S based managed services business model*

How does Uniphore's audio mining solution benefit my organization?

By combining voice capture with business intelligence, analytics and text mining provides valuable customer intelligence for marketing and competitive intelligence business functions. The benefits associated with implementing audio mining solution are:

• *Improved risk & liability management*

• *Real-time agent threshold monitoring alerts*

• *Intelligent and focus quality monitoring functionalities*

• *Equipped agents with critical data to secure sales*

• *Enhances customer experience with higher first contact resolution*

• *Advance Phonetics Engine Suited for Quick Response to Situations*

CREDIBILITY ///

What is the accuracy rate?

Accuracy levels depend upon the quality of the recording. Studio-based content will provide higher accuracy levels, but the system also provides a reasonable level of accuracy for telephone, public presentation and broadcast content.

The system recognizes "all words," not just keywords. The accuracy of preconfigured vocabularies can be further fine-tuned using the Vocabulary Tool to include organization-specific terms and proper names. This tool automatically customizes vocabularies with unique terms, such as industry-specific terminology or topics, resulting in outstanding recognition. The unsurpassed state-of-art-accuracy rate is consistently above 95%.

> " The unsurpassed state-of-art-accuracy rate is consistently above 95% "

Does Uniphore's audio mining support broadcast quality audio?

Yes, Uniphore's audio mining solution supports broadcast quality audio.

Does Uniphore's Audio Mining support video indexing?

No. However, it is possible to separate the audio portion from a video file, using free tools on the Internet, enabling our audio mining solution to index the corresponding audio content from a video file. The separation needs to be implemented as a separate process before invoking the audio mining indexer.

uniphore
Speech-based Mobility Solutions

What are the different voice loggers which can be integrated with Audio Mining?

Uniphore's audio mining solution can be integrated with voice loggers from service providers like NICE, VERINT and many more. Uniphore differs from other audio mining solution providers by integrating its audio mining solution with any of the voice loggers available in the market.

What are the different components of audio mining?

VOICE CAPTURE
The voice capture component improves transcription productivity by optimizing the digitized voice for speech processing, resulting in more accurate drafts.

SPEECH PROCESSING
The intelligent, background speech recognition engine interprets and formats audio.

TRANSCRIPTIONIST TOOLS
The transcriptionist tools are designed to increase productivity by speeding up the editing process.
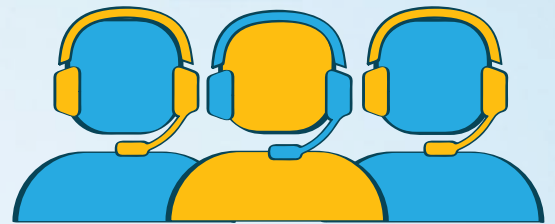
SAMPLE USE CASES ///

What are the industries that can use audio mining?

Audio mining plays a key-role in organizations where large volumes of audio data are collected from customer-service representative conversations. Most widely audio mining is used in customer contact centers, BPOs, Phone banking call centers, etc.

How is audio mining used in the contact center?

The audio mining solution is used in the contact centers to:

- *Train the workforce. For eg: Use "I don't know" as a key phrase to identify crucial knowledge gaps and develop targeted training.*

- *Product development and process improvement. For eg: If the words "bill" and "overcharge" start occurring together often, company can quickly course correct*

- *Corporate compliance: Actual matching of call scripts for monitoring of corporate compliance*

- *Risk identification. For eg: Identify the word 'Fraud' in the audio data to know why the customer or the service executive used the word and take necessary actions immediately.*

DIFFERENTIATORS ///

What are your key differentiators?

- *Better Context Analysis: We use sophisticated contextual information at the document, sentence, and phoneme level to produce accurate content and correct formatting.*

- *Better Recognition Models: We create language and formatting models specific to individual clients and to the work type being dictated.*

- *Better Transcription Integration: We incorporate specified formatting rules from each organization. Our speech recognition engine learns from new inputs, adapting the speech recognition models continually.*

uniphore
Speech-based Mobility Solutions

How does audio mining impact traditional transcription workflow?

Traditionally, an assigned transcriptionist produces a document by listening to a recorded dictation and typing the dictation using specialized transcription software and hardware. This is a labor-intensive – and therefore expensive – process.

State-of-the-art speech recognition technology in transcription workflow automates the transcription process, increasing transcription productivity and reducing transcription costs. When used in a transcription workflow, speech recognition is used to create a draft transcript of the recorded conversation. The transcriptionist listens to the dictation, verifies the draft, and makes corrections and further edits. Productivity gains are obtained because correcting recognition errors and editing text is faster than typing the entire dictation.

> " State-of-the-art speech recognition technology in transcription workflow automates the transcription process, increasing transcription productivityandreducing transcription costs "

What are the different modes to use audio mining in a contact center?

There are two modes in which the audio mining is used in a contact center:

• *Front-end speech recognition*

• *Back-end speech recognition*

un iphore
Speech-based Mobility Solutions

What is the difference between back-end and front-end speech recognition?

Front-end speech recognition takes place in real-time at the location where the contact center representative is conversing with the customer on a call.

On the other hand, back-end speech recognition takes place on a remote machine in a network. This is not real-time recognition; the results are produced in batch mode.

Back-end speech recognition used for transcription is inherently more difficult than speech recognition for interactive conversation (front-end speech recognition). One particular reason is that back-end recognition is typically tasked with transcribing speech that is recorded over the phone or a voice-recorder. Such recordings have higher noise levels and lower bandwidth than what is typical for recordings made with high-quality, noise-canceling, close-talking microphones. The recordings used for back-end transcription also generally reflect faster speaking rates and poorly-enunciated speech.

What are the audio formats that are supported by Uniphore's audio mining?

Uniphore's audio mining solution supports the following audio file types in both mono and stereo (8 kHz to 99 kHz):

- WAVE PCM
- MS ADPCM
- IMA ADPCM
- a-law
- mu-law
- VOX
- MP3
- WMA

ABOUT UNIPHORE ///

The ability to use speech to communicate is a primary reason for the evolutionary success of the human race. Uniphore's solutions extend this insight to the evolution of human-machine interaction. Uniphore's solutions allow any machine to understand and respond to natural human speech, thus enabling humans to use the most natural of communication modes, speech, to engage and instruct machines. Enterprises across industry, size and geographies deploy Uniphore's solution to dramatically improve employee productivity and deliver superior customer service.

As a leader of voice-based solutions, Uniphore has pioneered the development of mobile applications with the combined capabilities of Speech Recognition, Voice Biometrics, and Data. Uniphore boasts a roster of high-profile, satisfied customers across multiple verticals – Financial Service Providers (mobile commerce & banking), FMCGs & NBFCs (sales force automation), and Agriculture, Healthcare, & Education (content delivery services).

Since its inception in 2008, the company has grown at an exponential rate, and today it supports nearly half a million registered end users on its platforms every month. For more information on Uniphore visit www.uniphore.com.

Contact Us:

Uniphore Software Systems
India - +91 44 6646 9878
Dubai - +971 501528717
Philippines - +63 9278830228
Email: info@uniphore.com

uniphore
Speech-based Mobility Solutions