



ACTNow SDK 3.0

What is ACTNow?

ACTNow is an abbreviation for Audio Classification Technologies. ACTNow provides a complete set of technologies to extract different sorts of information from any audio or video file. Software developers can integrate the ACTNow technology by using the ACTNow SDK (Software Development Kit). ACTNow SDK is an easy-to-use Software Development Kit for creating professional software applications and solutions. ACTNow provides all core technologies for classifying and searching in audio or video content. Its integrated search functionalities will enable your software program to search for any particular audio segment or information in an audio file or stream. ACTNow is not only one technology but in fact it contains several "audio mining" technologies.

You can search extremely fast for words or phrases, sound clips, speech/music, specific speakers, if a certain audio track just starts or ends and many more.

How do I benefit?

By integrating ACTNow functionalities into your application you can basically analyze any Multimedia file such as audio for video files. Since ACTNow is a new technology it will make your application very unique.

Here are some examples how your application might benefit:

- If you develop a **recording application** (e.g. a **telephone or microphone or broadcast recorder**) you can make your audio or video files "searchable". You can offer to your customer the possibility of finding the information in his recordings he is actually looking for. This leads to huge timesavings compared to manual search. Imagine how long it would take to listen to all recordings just to find where something was spoken or a certain audio clip was played.
- If you develop a **content management** or **digital asset management** solution you will now be able to offer your customer a more

intelligent management solution. Your customers will be able to retrieve the content they are really looking for. With ACTNow you can extract completely automatically "new" information from your customer's content. Your customer's will be amazed how much information is actually in their multi-media content.

- Users of **E-Learning Applications** can locate the multimedia files that contain the requested information instantly. On top of that they can quickly navigate to the exact position in that audio file in an easy way.
- Developers of **telephone or broadcasting monitoring applications** can use ACTNow to spot words or jingles on the monitored channel online. This saves significant costs for your clients who currently monitor manually by listening.
- If you develop **security applications** and would like to verify the identity of a user you can simply use the ACTNow Speaker Detection feature to identify the speaker and the ACTNow Key Phrase detection technology to check if the person pronounces a given phrase in the same way.

ACTNow SDK Contains

- All core-technologies and parameter files as described
- C/C++ API to integrate all functionalities
- COM components for easy integration into any development environment (VB, C#, etc.)
- Online Help
- Sample applications
- Sample programs including source code



ACTNow Features

ACTNow Phonetic Index Search

The phonetic search engine enables you to search for words or phrases in any audio data. The search process comprises two passes. In the first pass feeding audio data to the API will create a phonetic index. Once an index is created you can store it on disk and retrieve it at any later time. Searching for any word or phrase in such index is extremely fast. To enable searches later on you can store and load such index. In a fraction of a second you can locate the exact position of a word in several hours of indexed sound. ACTNow supports

specific models for telephone recordings, microphone recordings and broadcast recordings to assure very high accuracy.

Currently supported languages:

- US English
- UK English
- German
- Hungarian
- Modern Standard Arabic

Please check for an updated list by sending a mail to sales@compure.com

ACTNow Speaker Detection

The speaker detection engine determines who is actually speaking. The speaker detection process comprises two passes. ACTNow needs to know "who" shall be identified later on. Therefore in the first pass the system will learn the speakers who shall be detected later on. In the second pass by feeding in audio data from a telephone, microphone or broadcasted stream ACTNow will detect which one (if any) of the previously trained speakers is actually speaking at particular parts of the audio stream.

ACTNow Audio Clip Detection

The audio clip detection engine finds the position of a previously specified clip in an audio stream. The search process comprises two passes. In the first pass you specify the audio clip (which can be a song, a jingle, a noise, or something similar...) by feeding the audio data to the API. Once a clip is specified you can detect this clip in any audio stream. The accuracy of this technology is extremely high even if the analyzed audio stream is noisy or distorted.

ACTNow Speech/Music Classification

The speech and music classification engine locates where music or speech can be found in an audio stream. Simply by feeding audio data from a telephone, microphone or radio/TV stream to the API you can locate the precise start and end times of each segment.

ACTNow Music Track Change Detection

The music track change detection engine tackles the difficult task to locate the exact position where a music track starts and ends in an audio stream. The accuracy depends on how distinct the different tracks are (e.g. in a "Mega Mix" it is very difficult to locate the changes).

ACTNow Key Phrase Detection

The key phrase detection searches for previously spoken words or phrases from the same speaker. The recognition process comprises two passes. In the first pass audio data from the spoken phrase are fed to the recognition engine to train the system on this phrase. In the second pass the actual audio stream to be analyzed is fed to the API. The recognition engine will mark where and with which likelihood the previously

trained phrase can be detected. This speech recognition technology is language independent.

ACTNow Multichannel Silence / Speech Detection

ACTNow automatically detects if on all audio streams either speech or silence is present. This is helpful to detect if e.g. all parties are silent or if 2 parties talk at the same time (one party interrupts the other).

System Requirements

ACTNow is designed to run on following Operating Systems:

- Microsoft ® Windows™ Server 2003
- Microsoft ® Windows ® XP
- Microsoft ® Windows ® 2000
- Microsoft ® Windows ® NT4.0
- Microsoft ® Windows 98 ®
- Microsoft ® Windows ME ®

Recommended Hardware Requirements:

- Pentium IV 500+ MHz
- 512MB RAM
- 20 GB Harddisk or higher

Evaluation Package

A complete evaluation package to test and evaluate our technology is available. Please contact sales@compure.com

Support

Compure provides experienced technical support services via e-mail, telephone or on-site consultancy. Please contact support@compure.com for further details.

Ordering and Contact

E-mail: sales@compure.com
Telephone: +36 1 3361 660
Fax: +36 1 3361 664
Web: www.compure.com

Due to continuous product improvements, specifications are subject to change without notice. Windows 98 SE, Windows Me, Windows 2000, Windows XP, Windows NT 4.0 are registered trademarks of Microsoft Corporation. Pentium is a registered trademark of Intel Corporation. Sound Blaster is a registered trademark of Creative Labs. All other trademarks or registered trademarks are the property of their respective companies. © 2000 - 2005 Compure Ltd, Hungary. All rights reserved