

Speaker  
recognition for  
stand-alone or  
Web applications

# VeriSpeak SDK



# VeriSpeak SDK

## Speaker recognition for stand-alone or Web applications

Document updated on **September 28, 2017**

### CONTENTS

VeriSpeak algorithm features and capabilities . . . . .	3
Contents of VeriSpeak Standard SDK and Extended SDK . . . . .	4
Biometric components description. . . . .	5
Basic recommendations for speaker recognition. . . . .	8
System requirements. . . . .	10
Technical specifications. . . . .	16
Reliability tests. . . . .	17
VeriSpeak Trial SDK and related products . . . . .	19
Licensing VeriSpeak SDK . . . . .	20
Prices for VeriSpeak products . . . . .	22

VeriSpeak voice identification technology is designed for biometric system developers and integrators. The text-dependent speaker recognition algorithm assures system security by checking both voice and phrase authenticity. Voiceprint templates can be matched in 1-to-1 (verification) and 1-to-many (identification) modes.

VeriSpeak is available as a software development kit that enables the development of stand-alone and Web-based applications on Microsoft Windows, Linux, Mac OS X, iOS and Android platforms.

- Text-dependent algorithm prevents unauthorized access with a covertly-recorded user voice.
- Two-factor authentication by checking voice biometrics and pass-phrase authenticity.
- Regular microphones and smartphones are suitable for recording user voices.
- Available as multiplatform SDK that supports multiple programming languages.
- Reasonable prices, flexible licensing and free customer support.



## VeriSpeak Algorithm Features and Capabilities

Neurotechnology's VeriSpeak is a **speaker recognition algorithm** designed for biometric system integrators. The VeriSpeak algorithm implements voice enrollment and voiceprint matching using proprietary sound processing technologies:

- **Text-dependent algorithm.** The text-dependent speaker recognition is based on saying the **same phrase for enrollment and verification**. The VeriSpeak algorithm determines if a voice sample matches the template that was extracted from a specific phrase. During enrollment, one or more phrases are requested from the person being enrolled. Later that person may be asked to pronounce a specific phrase for verification. This method assures protection against the use of a covertly recorded random phrase from that person.
- **Two-factor authentication with a passphrase.** The VeriSpeak voiceprint matching algorithm can be configured to work in a scenario where each user records a **unique phrase** (such as passphrase or an answer to a "secret question" that is **known only by the person** being enrolled). Later a person is recognized by his or her own specific phrase with a high degree of accuracy. The overall system security increases as both voice authenticity and passphrase are checked.
- **Text-independent algorithm.** The phrase-independent speaker recognition uses different phrases for user enrollment and recognition. This method is more convenient, as it does not require each user to remember the passphrase. It may be combined with the text-dependent algorithm to perform faster text-independent search with further phrase verification using the more reliable text-dependent algorithm.
- **Automatic voice activity detection.** VeriSpeak is able to detect when users start and finish speaking.
- **Liveness detection.** A system may request each user to enroll a set of unique phrases. Later the user will be requested to say a specific phrase from the enrolled set. This way the system can ensure that a live person is being verified (as opposed to an impostor who uses a voice recording).
- **Identification capability.** VeriSpeak functions can be used in 1-to-1 matching (verification) and **1-to-many** (identification) modes.
- **Multiple samples of the same phrase.** A template may store several voice records with the same phrase to improve recognition reliability. Certain natural voice variations (i.e. hoarse voice) or environment changes (i.e. office and outdoors) can be stored in the same template.
- **Fused matching.** A system may ask users to pronounce several specific phrases during speaker verification or identification and match each audio sample against records in the database. The VeriSpeak algorithm can fuse the matching results for each phrase together to improve matching reliability.



## Contents of VeriSpeak Standard SDK and Extended SDK

VeriSpeak SDK is based on VeriSpeak voice recognition technology and is designed for biometric systems developers and integrators. The SDK allows rapid development of biometric applications using functions from the VeriSpeak algorithm. VeriSpeak can be easily integrated into the customer's security system. The integrator has complete control over SDK data input and output.

VeriSpeak is available as the following SDKs:

- **VeriSpeak Standard SDK** is designed for PC-based, embedded or mobile biometric application development. It includes Voice Matcher and Extractor component licenses, programming samples and tutorials and software documentation. The SDK enables the development of biometric applications for Microsoft Windows, Linux, Mac OS X and Android operating systems.
- **VeriSpeak Extended SDK** is designed for biometric **Web-based** and network application development. It includes all features and components of the Standard SDK with the addition of Voice Client component licenses for PCs and Android devices, sample client applications, tutorials and a **ready-to-use matching server** component.

The table below compares VeriSpeak 10.0 Standard SDK and VeriSpeak 10.0 Extended SDK. See the licensing model for more information on specific license types.

Component licenses that are included with a specific SDK		
	VeriSpeak 10.0 Standard SDK	VeriSpeak 10.0 Extended SDK
• Voice Matcher	1 single computer license	1 single computer license
• Embedded Voice Matcher for Android	1 single computer license	1 single computer license
• Embedded Voice Matcher for iOS	1 single computer license	1 single computer license
• Embedded Voice Matcher for ARM Linux	1 single computer license	1 single computer license
• Voice Client		3 single computer licenses
• Embedded Voice Client for Android		3 single computer licenses
• Embedded Voice Client for iOS		3 single computer licenses
• Embedded Voice Client for ARM Linux		3 single computer licenses
• Voice Extractor	1 single computer license	1 single computer license
• Embedded Voice Extractor for Android	1 single computer license	1 single computer license
• Embedded Voice Extractor for iOS	1 single computer license	1 single computer license
• Embedded Voice Extractor for ARM Linux	1 single computer license	1 single computer license
• Matching Server		+

VeriSpeak SDK includes programming samples and tutorials that show how to use the components of the SDK to perform voice template extraction or matching against other templates. The samples and tutorials are available for these programming languages and platforms:

	Windows	Linux	Mac OS X	iOS	Android
<b>Programming samples</b>					
• C/C++	+	+	+		
• Objective-C				+	
• C#	+				
• Java	+	+	+		+
• Visual Basic .NET	+				
<b>Programming tutorials</b>					
• C	+	+	+		
• C#	+				
• Visual Basic .NET	+				
• Java	+	+	+		+



## Biometric Components Description

### Voice Matcher

The Voice Matcher performs voice template matching in 1-to-1 (verification) and 1-to-many (identification) modes on PC and Mac platforms. Also the Voice Matcher component includes fused matching algorithm that allows to increase template matching reliability by matching templates that contain fingerprint, face, voice and/or iris records (note that matching fingerprints, irises and faces requires to purchase Fingerprint Matcher, Iris Matcher and Face Matcher components correspondingly - these components are available in *VeriFinger 10.0 SDK*, *VeriEye 10.0 SDK* and *VeriLook 10.0 SDK* respectively; see these products brochures for more information).

The Voice Matcher component matches **8,000 voiceprints per second** and is designed to be used in **desktop** or mobile biometric systems, which run on PCs or laptops with at least Intel **Core i7-4771** (3.5 GHz) processor.

One Voice Matcher license is included with VeriSpeak 10.0 Standard SDK and VeriSpeak 10.0 Extended SDK. More licenses for this component can be purchased any time by VeriSpeak 10.0 SDK customers.

### Embedded Voice Matcher

The Embedded Voice Matcher has the same functionality as the Voice Matcher. It matches **100 voiceprints per second** and is designed to be used in **embedded** or **mobile** biometric systems, which run on ARM Linux, **Android** or **iOS** devices. The Android devices should be based on at least **Snapdragon S4** system-on-chip (**Krait 300** processor with 4 cores running at 1.51 GHz).

One Embedded Voice Matcher license for each of Android, iOS and ARM Linux platforms is included with VeriSpeak 10.0 Standard SDK and VeriSpeak 10.0 Extended SDK. More licenses for this component can be purchased any time by VeriSpeak 10.0 SDK customers.

### Voice Client

The Voice Client component currently has the same functionality as Voice Extractor component. It is intended for using on PC- and Mac-based clients of web-based and network-based biometric systems.

The Voice Client extracts a single voiceprint template in **0.6 seconds**. The specified performance requires a **PC or laptop** with at least Intel **Core 2 Q9400** (2.67 GHz) processor.

Three licenses for the Voice Client component are included with VeriSpeak 10.0 Extended SDK. More licenses for this component can be purchased any time by VeriSpeak 10.0 Extended SDK customers.

### Embedded Voice Client

The Embedded Voice Client component currently has the same functionality as Embedded Voice Extractor component. It is intended to be used in **embedded** or **mobile** biometric systems, which run on ARM Linux, **Android** or **iOS** devices. The Android devices should be based on at least **Snapdragon S4** system-on-chip (**Krait 300** processor with 4 cores running at 1.51 GHz).

Three Embedded Voice Client licenses for each of Android, iOS and ARM Linux platforms are included with VeriSpeak 10.0 Extended SDK. More licenses for this component can be purchased any time by VeriSpeak 10.0 Extended SDK customers.



## Voice Extractor

Voice Extractor creates voice templates from audio samples on PC and Mac platform.

See technical specifications the size of voice template and the requirements for voice record.

The component extracts a single voiceprint template in **1.34 seconds**. The specified performance requires a **PC or laptop** with at least Intel **Core 2 Q9400** (2.67 GHz) processor.

One Voice Extractor license is included with VeriSpeak 10.0 Standard SDK and VeriSpeak 10.0 Extended SDK. More licenses for this component can be purchased any time by VeriSpeak 10.0 SDK customers.

## Embedded Voice Extractor

The Embedded Voice Extractor has the same functionality as the Voice Extractor and is designed to be run on ARM Linux, **Android** or **iOS** devices. The Android devices should be based on at least **Snapdragon S4** system-on-chip (**Krait 300** processor with 4 cores running at 1.51 GHz). The component extracts a single voiceprint template in **1.34 seconds**.

One Embedded Voice Extractor license for each of Android, iOS and ARM Linux platforms is included with VeriSpeak 10.0 Standard SDK and VeriSpeak 10.0 Extended SDK. More licenses for this component can be purchased any time by VeriSpeak 10.0 SDK customers.



## Matching Server

The Matching Server is ready-to-use software intended for building moderate size web-based and other network-based systems like local single- or multi-biometric identification system. The Server software runs on a server PC and allows to perform the biometric template matching on server side using Voice Matcher component.

**Multi-biometric matching** can be enabled by running components for fingerprint, face, voiceprint and iris matching on the same machine.

**Client communication module** that allows sending a task to the Matching Server, querying status of the task, getting the results and removing the task from server, is included with MegaMatcher 10.0 SDK, VeriFinger 10.0 SDK, VeriLook 10.0 SDK, VeriSpeak 10.0 SDK and VeriEye 10.0 SDK. This module hides all low level communications and provides high-level API for the developer.

The components and database support modules with source codes included for Matching Server component are listed in the table below. Custom modules for working with other databases can also be developed by integrator and used with the Matching Server software.

The table below shows what components are available with Matching Server software.

Components	Microsoft Windows 32 & 64 bit	Linux 32 & 64 bit	Mac OS X
• Matching server software	+	+	+
• Server administration tool API	+	+	
<b>Database support modules</b>			
• Microsoft SQL Server	+		
• PostgreSQL	+	+	
• MySQL	+	+	
• Oracle	+	+	
• SQLite	+	+	+
<b>Programming samples</b>			
• C# client	+		
• Visual Basic .NET client	+		
• Java web client	+	+	+
<b>Programming tutorials</b>			
• C/C++	+	+	
• C#	+		
• Visual Basic .NET	+		
• Delphi	+		

The Matching Server component requires a **special license** that allows to run the component on all machines that run the fingerprint, face, iris or palm print matching components obtained by an integrator. The Matching Server software is included with VeriSpeak Extended SDK.

Also the Matching Server component is included with these Neurotechnology SDKs (see their brochures for more info):

- MegaMatcher 10.0 Standard or MegaMatcher 10.0 Extended SDK;
- VeriFinger 10.0 Extended SDK;
- VeriLook 10.0 Extended SDK;
- VeriEye 10.0 Extended SDK.



## Basic Recommendations for Speaker Recognition

The speaker recognition accuracy of VeriSpeak depends on the audio quality during enrollment and identification. Certain constraints should be noted before or during algorithm integration into a speaker recognition system. Other variables may be overcome by enrollment with the same phrase in different environments.

**Voice samples of at least 2 seconds in length are recommended** to assure recognition quality.

### General Security

A **passphrase should be kept secret and not spoken in an environment where others may hear it** if the speaker recognition system is used in a scenario with unique phrases for each user.

The **text-independent** speaker recognition may be **vulnerable** to attack with a **covertly recorded phrase** from a person. **Passphrase verification** or **two-factor authentication** (i.e. requirement to type a password) will increase the overall system **security**.

### Microphones

There are no particular constraints on models or manufacturers when using regular PC microphones, headsets or built-in laptop microphones. However these factors should be noted:

- The **same microphone model** is recommended (if possible) for use during both enrollment and recognition, as different models may produce different sound quality. Some models may also introduce specific noise or distortion into the audio, or may include certain hardware sound processing, which will not be present when using a different model. This is also the recommended procedure when using **smartphones** or **tablets**, as different device models may alter the recording of the voice in different ways.
- The same **microphone position** and distance is recommended during enrollment and recognition. **Headsets** provide optimal distance between user and microphone; this distance is recommended when non-headset microphones are used.
- **Web cam built-in** microphones should be **used with care**, as they are usually positioned at a rather long distance from the user and may provide lower sound quality. The sound quality may be affected if users subsequently change their position relative to the web cam.

### Sound Settings

**Settings for clear sound** must be ensured; some audio software, hardware or drivers may have sound modification **enabled by default**. For example, the Microsoft Windows OS usually has, by default, sound boost enabled.

A **minimum 11025 Hz** sampling rate, with at least **16-bit** depth, should be used during voice recording.





## Environment Constraints

The VeriSpeak speaker recognition algorithm is sensitive to **noise** or **loud voices** in the **background**; they may **interfere** with the user's voice and affect the recognition results. These solutions may be considered to reduce or eliminate these problems:

- A **quiet environment** for enrollment and recognition.
- **Several samples of the same phrase** recorded in different environments can be stored in a biometric template. Later the user will be matched against these samples with much higher recognition quality.
- **Close-range microphones** (like those in headsets) that are not affected by distant sources of sound.
- Third-party or custom solutions for background noise reduction, such as using two separate microphones for recording user voice and background sound, and later subtracting the background noise from the recording.

## User Behavior and Voice Changes

Natural voice changes may affect speaker recognition accuracy:

- A temporarily **hoarse voice** caused by a cold or other sickness
- Different **emotional states** that affect voice (i.e. a cheerful voice versus a tired voice)
- Different **pronunciation speeds** during enrollment and identification

The aforementioned voice and user behavior changes can be managed in two ways:

- **Separate enrollments** for the altered voice, storing the records to the same person's template;
- A **controlled, neutral voice** during enrollment and identification.



## System requirements

There are specific requirements for each platform which will run VeriSpeak-based applications.

### Microsoft Windows platform requirements

- Microsoft Windows **Vista / 7 / 8 / 10**, 32-bit or 64-bit.
  - Windows XP is no longer supported in this version of the SDK. If your product requires to support Windows XP, you may consider the previous version of the SDK. Please contact us for more information.
- PC or laptop with **x86 (32-bit)** or **x86-64 (64-bit)** compatible processors.
  - 2 GHz or better processor is recommended.
  - **SSE2 support is required.** Processors that do not support SSE2 cannot run the VeriSpeak algorithm. Please check if a particular processor model supports SSE2 instruction set.
- At least **128 MB of free RAM** should be available for the application. Additional RAM is required for applications that perform 1-to-many identification, as all biometric templates need to be stored in RAM for matching. For example, **1,000 templates** (each with 1 voiceprint record) require about **5 MB of additional RAM**.
- **Free space on hard disk drive (HDD):**
  - at least 1 GB required for the development.
  - 100 MB required for VeriSpeak components deployment.
  - Additional space would be required in these cases:
    - VeriSpeak does not require the original voice sample to be stored for the matching; only the templates need to be stored. However, storing voice samples on hard drive for the potential future usage is recommended.
    - Usually a database engine runs on a separate computer (back-end server). However, DB engine can be installed on the same computer for standalone applications. In this case HDD space for templates storage must be available. For example, 1,000 templates (each with 1 voiceprint record) stored using a relational database would require about 5 MB of free HDD space.
- **Microphone.** Any microphone that is supported by the operating system can be used.
- **Database engine** or connection with it. VeriSpeak templates can be saved into any DB (including files) supporting binary data saving. VeriSpeak Extended SDK contains the following support modules for Matching Server on Microsoft Windows platform: **Microsoft SQL Server, MySQL, Oracle, PostgreSQL** and **SQLite**.
- **Network/LAN connection (TCP/IP)** for client/server applications. Also, network connection is required for using Matching server component (included in VeriSpeak Extended SDK). Communication with Matching server is not encrypted, therefore, if communication must be secured, a dedicated network (not accessible outside the system) or a secured network (such as VPN; VPN must be configured using operating system or third party tools) is recommended.
- Microsoft **.NET framework 3.5** or newer (for .NET components usage).
- One of following **development environments** for application development:
  - Microsoft Visual Studio 2012 or newer (for application development under C/C++, C#, Visual Basic .Net)
  - Sun Java 1.6 SDK or later.



## Android platform requirements

- A smartphone or tablet that is running **Android 4.4 (API level 19)** OS or newer.
  - API level 22 is the recommended **target** for code compilation.
  - If you have a custom Android-based device or development board, contact us to find out if it is supported.
- ARM-based **1.5 GHz processor recommended** for voiceprint processing in the specified time. Slower processors may be also used, but the voiceprint processing will take longer time.
- At least **20 MB of free RAM** should be available for the application. Additional RAM is required for applications that perform 1-to-many identification, as all biometric templates need to be stored in RAM for matching. For example, **1,000 templates** (each with 1 voiceprint record) require about **5 MB of additional RAM**.
- **Free storage** space (built-in flash or external memory card):
  - 30 MB required for embedded voice components deployment for each separate application.
  - Additional space would be required if an application needs to store original voice samples. VeriSpeak does not require the original voice sample to be stored for the matching; only the templates need to be stored.
- Any smartphone's or tablet's **built-in or headset microphone** which is supported by Android OS.
- **Network/LAN connection (TCP/IP)** for client/server applications. Also, network connection is required for using Matching server component (included in VeriSpeak Extended SDK). Communication with Matching server is not encrypted, therefore, if communication must be secured, a dedicated network (not accessible outside the system) or a secured network (such as VPN; VPN must be configured using operating system or third party tools) is recommended.
- **PC-side development** environment requirements:
  - Java SE JDK 6 (or higher)
  - Eclipse Indigo (3.7) IDE
  - Android development environment (at least API level 19 required)
  - One of the following build automation systems:
    - Apache Maven 3.1.x or newer
    - Gradle 2.10 or newer
  - Internet connection for activating VeriSpeak component licenses



## iOS platform requirements

- One of the following devices, running **iOS 8.0** or newer:
  - **iPhone 5** or newer iPhone.
  - **iPad 2** or newer iPad, including iPad Mini and iPad Air models.
  - **iPod Touch 6th Generation** or newer iPod.
- At least **20 MB of free RAM** should be available for the application. Additional RAM is required for applications that perform 1-to-many identification, as all biometric templates need to be stored in RAM for matching. For example, **1,000 templates** (each with 1 voiceprint record) require about **5 MB of additional RAM**.
- **Free storage** space (built-in flash or external memory card):
  - 30 MB required for embedded voice components deployment for each separate application.
  - Additional space would be required if an application needs to store original voice samples. VeriSpeak does not require the original voice sample to be stored for the matching; only the templates need to be stored.
- Any smartphone's or tablet's **built-in or headset microphone** which is supported by iOS.
- **Network/LAN connection (TCP/IP)** for client/server applications. Also, network connection is required for using Matching server component (included in VeriSpeak Extended SDK). Communication with Matching server is not encrypted, therefore, if communication must be secured, a dedicated network (not accessible outside the system) or a secured network (such as VPN; VPN must be configured using operating system or third party tools) is recommended.
- **Development environment** requirements:
  - a Mac running Mac OS X 10.10.x or newer.
  - Xcode 6.4 or newer.



## Mac OS X platform requirements

- A Mac running **Mac OS X 10.7.x** or newer. 2 GHz or better processor is recommended.
- At least **128 MB of free RAM** should be available for the application. Additional RAM is required for applications that perform 1-to-many identification, as all biometric templates need to be stored in RAM for matching. For example, **1,000 templates** (each with 1 voiceprint record) require about **5 MB of additional RAM**.
- **Free space on hard disk drive (HDD):**
  - at least 1 GB required for the development.
  - 100 MB required for VeriSpeak components deployment.
  - Additional space would be required in these cases:
    - VeriSpeak does not require the original voice sample to be stored for the matching; only the templates need to be stored. However, storing voice samples on hard drive for the potential future usage is recommended.
    - Usually a database engine runs on a separate computer (back-end server). However, DB engine can be installed on the same computer for standalone applications. In this case HDD space for templates storage must be available. For example, 1,000 templates (each with 1 voiceprint record) stored using a relational database would require about 5 MB of free HDD space. Also, the database engine itself requires HDD space for running. Please refer to HDD space requirements from the database engine providers.
- **Microphone.** Any microphone that is supported by the operating system can be used.
- **Database engine** or connection with it. VeriSpeak templates can be saved into any DB (including files) supporting binary data saving. VeriSpeak Extended SDK contains **SQLite** support modules for Matching Server on Mac OS X platform.
- **Network/LAN connection (TCP/IP)** for client/server applications. Also, network connection is required for using Matching server component (included in VeriSpeak Extended SDK). Communication with Matching server is not encrypted, therefore, if communication must be secured, a dedicated network (not accessible outside the system) or a secured network (such as VPN; VPN must be configured using operating system or third party tools) is recommended.
- Specific requirements for **application development:**
  - XCode 4.3 or newer
  - wxWidgets 3.0.0 or newer libs and dev packages (to build and run SDK samples and applications based on them)
  - Qt 4.8 or newer libs, dev and qmake packages (to build and run SDK samples and applications based on them)
  - GNU Make 3.81 or newer (to build samples and tutorials development)
  - Sun Java 1.6 SDK or later.



## Linux x86 / x86-64 platform requirements

- Linux 2.6 or newer kernel (32-bit or 64-bit) is required. **Linux 3.0 kernel** or newer is recommended.
- PC or laptop with **x86 (32-bit)** or **x86-64 (64-bit)** compatible processors.
  - 2 GHz or better processor is recommended.
  - **SSE2 support is required.** Processors that do not support SSE2 cannot run the VeriSpeak algorithm. Please check if a particular processor model supports SSE2 instruction set.
- At least **128 MB of free RAM** should be available for the application. Additional RAM is required for applications that perform 1-to-many identification, as all biometric templates need to be stored in RAM for matching. For example, **1,000 templates** (each with 1 voiceprint record) require about **5 MB of additional RAM**.
- **Free space on hard disk drive (HDD):**
  - at least 1 GB required for the development.
  - 100 MB required for VeriSpeak components deployment.
  - Additional space would be required in these cases:
    - VeriSpeak does not require the original voice sample to be stored for the matching; only the templates need to be stored. However, storing voice samples on hard drive for the potential future usage is recommended.
    - Usually a database engine runs on a separate computer (back-end server). However, DB engine can be installed on the same computer for standalone applications. In this case HDD space for templates storage must be available. For example, 10,000 templates (each with 1 voiceprint record) stored using a relational database would require about 50 MB of free HDD space. Also, the database engine itself requires HDD space for running. Please refer to HDD space requirements from the database engine providers.
- **Microphone.** Any microphone that is supported by the operating system can be used.
- glibc 2.11.3 library or newer
- libasound 1.0.x or newer (for voice capture)
- libgudev-1.0 164-3 or newer (for microphone usage)
- **Database engine** or connection with it. VeriSpeak templates can be saved into any DB (including files) supporting binary data saving. VeriSpeak Extended SDK contains **MySQL, Oracle, PostgreSQL** and **SQLite** support modules for Matching Server on Linux x86 / x86-64 platforms.
- **Network/LAN connection (TCP/IP)** for client/server applications. Also, network connection is required for using Matching server component (included in VeriSpeak Extended SDK). Communication with Matching server is not encrypted, therefore, if communication must be secured, a dedicated network (not accessible outside the system) or a secured network (such as VPN; VPN must be configured using operating system or third party tools) is recommended.
- Specific requirements for **application development:**
  - wxWidgets 3.0.0 or newer libs and dev packages (to build and run SDK samples and applications based on them)
  - Qt 4.8 or newer libs, dev and qmake packages (to build and run SDK samples and applications based on them)
  - GCC-4.4.x or newer
  - GNU Make 3.81 or newer (to build samples and tutorials development)
  - Sun Java 1.6 SDK or later.
  - pkg-config-0.21 or newer (optional; only for Matching Server database support modules compilation)



## ARM Linux platform requirements

We recommend to contact us and report the specifications of a target device to find out if it will be suitable for running VeriSpeak-based applications. There is a list of common requirements for ARM Linux platform:

- A device with ARM-based processor, running **Linux 3.2 kernel** or newer.
- ARM-based **1.5 GHz processor recommended** for voiceprint processing in the specified time. .
  - **ARMHF architecture (EABI 32-bit hard-float ARMv7)** is required.
  - Lower clock-rate processors may be also used, but the voiceprint processing will take longer time
- At least **20 MB of free RAM** should be available for the application. Additional RAM is required for applications that perform 1-to-many identification, as all biometric templates need to be stored in RAM for matching. For example, **1,000 templates** (each with 1 voiceprint record) require about **5 MB of additional RAM**.
- **Free storage** space (built-in flash or external memory card):
  - 100 MB required for VeriSpeak components deployment.
  - Additional space would be required if an application needs to store original voice samples. VeriSpeak does not require the original voice sample to be stored for the matching; only the templates need to be stored.
- **Microphone.** Any microphone that is supported by the operating system can be used.
- glibc 2.13 library or newer
- libasound 1.0.x or newer (for voice capture)
- libgudev-1.0 164-3 or newer (for microphone usage)
- libstdc++-v3 4.7.2 or newer.
- **Network/LAN connection (TCP/IP)** for client/server applications. Also, network connection is required for using Matching server component (included in VeriSpeak Extended SDK). Communication with Matching server is not encrypted, therefore, if communication must be secured, a dedicated network (not accessible outside the system) or a secured network (such as VPN; VPN must be configured using operating system or third party tools) is recommended.
- **Development environment** requirements:
  - GCC-4.4.x or newer
  - GNU Make 3.81 or newer
  - JDK 1.6 or later



## Technical Specifications

A minimum 11025 Hz sampling rate, with at least 16-bit depth, should be used during voice recording.

Voice samples of at least 2 seconds in length are recommended to assure recognition quality. Longer voice samples will further improve recognition.

See above the list of recommendations and constraints for speaker recognition.

All voice templates should be loaded into RAM before identification, thus the maximum voice template database size is limited by the amount of available RAM.

The voiceprint template size has **linear dependence** on the voice sample length. For example, when using voice samples that are 2 times shorter, the template size values will be 2 times smaller.

VeriSpeak 10.0 text-dependent engine can perform template matching in two modes:

- **Fixed phrase** – each subject in the database has recorded the same phrase. This mode provides **faster matching**, but lower reliability.
- **Unique phrase** – each subject in the database has recorded a unique phrase. This mode provides **higher reliability**, but slower matching.

VeriSpeak biometric template extraction and matching algorithm is designed to run on **multi-core processors** allowing to reach maximum possible performance on the used hardware.

VeriSpeak 10.0 text-dependent algorithm specifications				
	Android-based platform		PC-based platform	
Template extraction components	Embedded Voice Extractor	Embedded Voice Client	Voice Extractor	Voice Client
Template extraction time (seconds)	1.34 <sup>(1)</sup>	1.20 <sup>(1)</sup>	1.34 <sup>(2)</sup>	0.60 <sup>(2)</sup>
Template matching components	Embedded Voice Matcher		Voice Matcher	
Template matching speed fixed phrase mode (voiceprints per second)	100 <sup>(1)</sup>		8,000 <sup>(3)</sup>	
Template matching speed unique phrase mode (voiceprints per second)	20 <sup>(1)</sup>		1,700 <sup>(3)</sup>	
Single voiceprint record size in a template, when 5 second long voice samples used (bytes)	3,500 - 4,500			

Notes:

(1) Requires to be run on Android devices based on at least Snapdragon S4 system-on-chip with Krait 300 processor (4 cores, 1.51 GHz).

(2) Requires to be run on PC or laptop with at least Intel Core 2 Q9400 quad-core processor (2.67 GHz) to reach the specified performance.

(3) Requires to be run on PC or laptop with at least Intel Core i7-4771 quad-core processor (3.5 GHz) to reach the specified performance.





## Reliability Tests

The VeriSpeak 10.0 algorithm has been tested with voice samples taken from the XM2VTS Database, as well as with voice voice from Neurotechnology internal database.

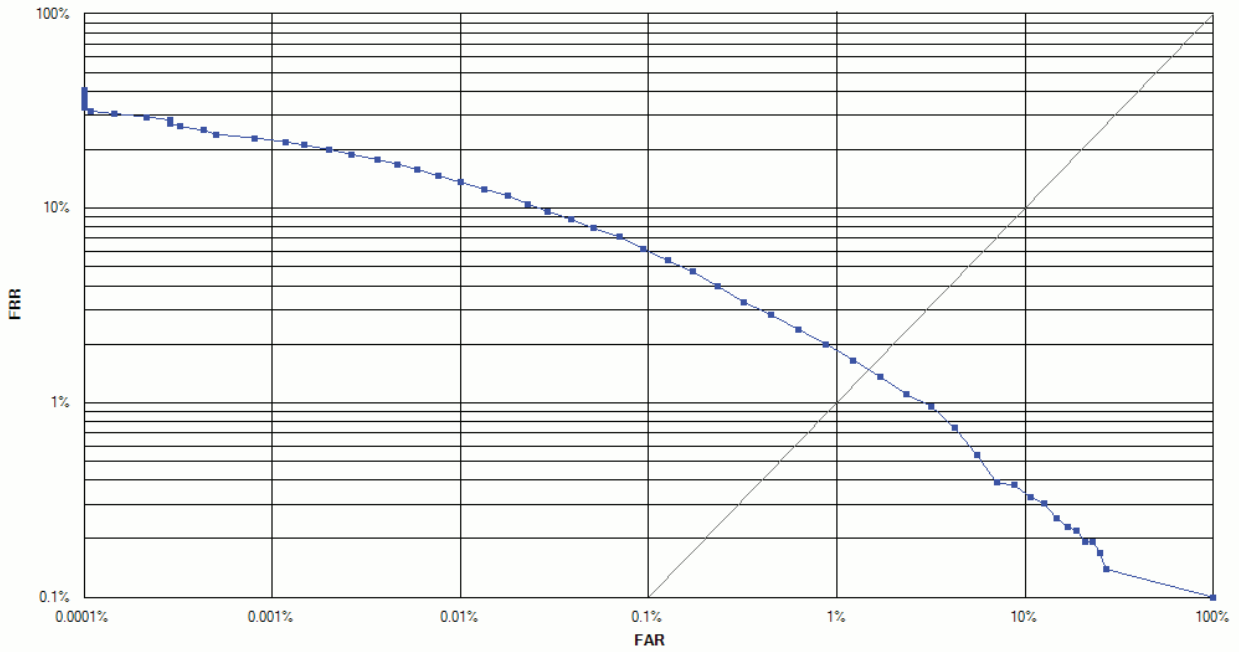
More information on XM2VTS database is available at <http://www.ee.surrey.ac.uk/CVSSP/xm2vtsdb/>

These voice template extraction and matching experiments were performed:

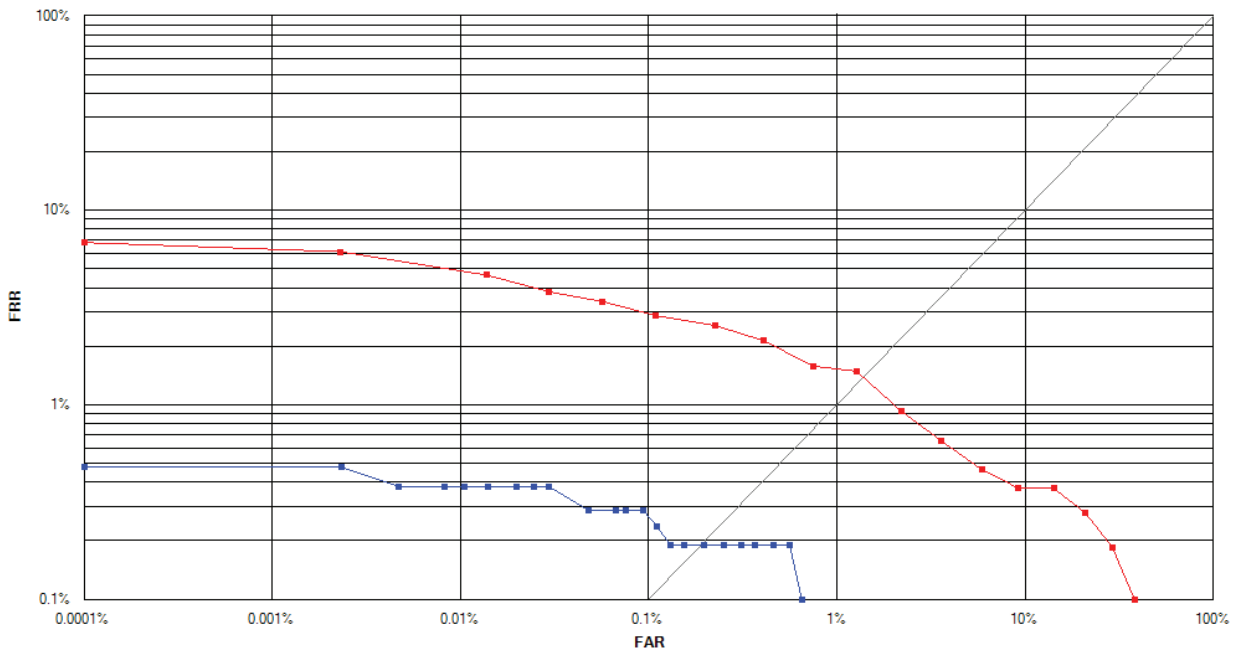
- **Experiment 1** used voice samples from the **XM2VTS** database. All samples include the same **fixed phrase** pronounced by all subjects.
- **Experiment 2** used voice samples from Neurotechnology internal voice database 1. All samples include the same **fixed phrase** pronounced by all subjects.
- **Experiment 3** used voice samples from Neurotechnology internal voice database 2. Each subject pronounced a **unique phrase** during his/her recording.

VeriSpeak 10.0 text-dependent algorithm tests with XM2VTS and Neurotechnology internal databases			
	Experiment 1	Experiment 2	Experiment 3
Total voice samples in the database	2360	309	305
Subjects in the database	295	42	42
Recording sessions per subject	8	1 - 10	1 - 10
Average voice sample length (seconds)	6.167	4.975	6.214
FRR at 0.1 % FAR	6.180 %	3.386 %	0.285 %

Receiver operation characteristics (**ROC**) curves are usually used to demonstrate the recognition quality of an algorithm. ROC curves show the dependence of false rejection rate (**FRR**) on the false acceptance rate (**FAR**). Charts with ROC curves for each of the experiments are available on the next page.



VeriSpeak 10.0 SDK text-dependent matching engine with voice records from XM2VTS database; all persons say the same phrase.



VeriSpeak 10.0 SDK text-dependent matching engine with voice records from Neurotechnology internal database:  
 ■ all persons say the same phrase  
 ■ each person says a unique phrase



## VeriSpeak Trial SDK and Related Products

VeriSpeak 30-day SDK Trial is available for downloading at [www.neurotechnology.com/download.html](http://www.neurotechnology.com/download.html).

These products are related to VeriSpeak SDK:

- **MegaMatcher SDK** – for development of AFIS or multi-biometric face, fingerprint, voice, iris and palm print identification products. See “MegaMatcher SDK” brochure for more information.
- **NPointer** – a freeware application for voice-based and gestural computer control. The application accepts voice commands for actions like clicking, scrolling and dragging. Also it captures hand movements with a webcam and converts them into pointer movements.  
NPointer can be downloaded at: <http://download.neurotechnology.com/NPointer.zip>



## Licensing VeriSpeak SDK

### Product Development

An integrator should obtain either a VeriSpeak 10.0 Standard SDK (EUR 339) or VeriSpeak 10.0 Extended SDK (EUR 859) to develop a end-user product based on VeriSpeak technology. The SDK needs to be purchased just once and may be used for all projects and by all the developers within the integrator's company.

See the "Contents of VeriSpeak Standard SDK and Extended SDK" chapter (page 4) for the list of component licenses included with VeriSpeak 10.0 Standard and VeriSpeak 10.0 Extended SDK.

Integrators can obtain additional component licenses if more component licenses are required for the development process.

### Product Deployment

To deploy their developed products, an integrator need obtain licenses of components for every **computer or device**, where component will be installed together with integrator's product. See Product Advisor to find out what specific components will be needed for the deployment of your system. Integrators can purchase additional VeriSpeak component licenses if required at anytime.

### License activation options

The components are copy-protected. The following license activation options are available:

- **Serial numbers** are used to activate licenses for particular VeriSpeak components on particular computer or device. The activation is done via the Internet or by email. After activation the network connection is not required for single computer license usage.
 

Notes:

  1. Activation by serial number is **not suitable for iOS and ARM-Linux** platforms, except BeagleBone Black and Raspberry Pi 3 devices..
  2. Activation by serial number is **not suitable for virtual environments**.
- **Internet activation.** A special **license file** is stored on a computer or a mobile or embedded device; the license file allows to run particular VeriSpeak components on that computer or device after **checking** the license over the Internet. **Internet connection** should be available periodically for a short amount of time. A single computer license can be **transferred** to another computer or device by moving the license file there and waiting until the previous activation expires.
- **Volume License Manager.** Licenses may be stored in a volume license manager **dongle**. License activation using volume license manager may be performed without connection to the Internet and is suitable for virtual environments. Volume license manager is **used on site by integrators or end users** to manage licenses for VeriSpeak components in the following ways:
  1. **Activating single computer licenses** – An installation license for a VeriSpeak component will be activated for use on a particular computer. The number of available licenses in the license manager will be decreased by the number of activated licenses. .
  2. **Managing single computer licenses via a LAN or the Internet** – The license manager allows the management of installation licenses for VeriSpeak components across multiple computers or mobile/embedded devices in a LAN or over the Internet. The number of managed licenses is limited by the number of licenses in the license manager. No license activation is required and the license quantity is not decreased. Once issued, the license is assigned to a specific computer or device on the network.
  3. **Using license manager as a dongle** – A volume license manager containing at least one license for a VeriSpeak component may be used as a dongle, allowing the VeriSpeak component to run on the particular computer where the dongle is attached.



## Licenses Validity

All SDK and component licenses are perpetual and do not have expiration. There are no annual fee or any other fees except license purchasing fee. It is possible to move licenses from one computer or device to another. Neurotechnology provides a way to renew the license if the computer undergoes changes due to technical maintenance.

## Licensing Agreement

The Licensing Agreement ([http://neurotechnology.com/mm\\_100\\_sla.html](http://neurotechnology.com/mm_100_sla.html)) contains all licensing terms and conditions.

Note that you unambiguously accept this agreement by placing an order using Neurotechnology online ordering service or by email or other means of communications. Please read the agreement before making an order.

## Other licensing options

- **VAR License.** The above described licensing model is intended for **end-user** product developers. Integrators who want to develop and sell a VeriSpeak-based development tool (with API, programming possibilities, programming samples, etc.), must obtain permission from Neurotechnology and **sign** a special VAR agreement. For more information please contact us.
- **Enterprise License.** The VeriSpeak enterprise license allows an **unlimited use** of VeriSpeak components in end-user products for a specific territory, market segment or project. Specific restrictions would be included in the licensing agreement. The enterprise license price depends on the application size and the number of potential users of the application within the designated territory, market segment or project. For more information please contact us.



## Prices for VeriSpeak products

- The prices are **effective June 21, 2017**. The prices may change in the future, so please **download and review the latest version** of the brochure before making an order.
- Quantity discounts do not accumulate over time.
- Prices do not include local import duties or taxes.
- Product shipping costs depend on delivery country.
- Customers with Solution Partner status are eligible for product discounts.

VeriSpeak SDK			
VeriSpeak 10.0 Standard SDK			€ 339.00
VeriSpeak 10.0 Extended SDK			€ 859.00
Voice Client concurrent licenses			
Price per license			€ 470.00
Voice components for PCs (prices per single computer license)			
Quantity	Voice Extractor	Voice Client <sup>(1)</sup>	Voice Matcher
1-9	€ 20.00	€ 20.00	€ 25.00
10-19	€ 15.00	€ 15.00	€ 18.00
20-49	€ 13.00	€ 13.00	€ 16.00
50-99	€ 11.00	€ 11.00	€ 14.00
100-199	€ 10.00	€ 10.00	€ 12.50
200-499	€ 9.00	€ 9.00	€ 11.00
500-999	€ 8.00	€ 8.00	€ 10.00
1000-1999	€ 7.00	€ 7.00	€ 9.00
2000-3999	€ 6.40	€ 6.40	€ 8.00
4000-7999	€ 5.80	€ 5.80	€ 7.00
8000 and more	Please contact us for more information		
Embedded voice components for Android devices (prices per single computer license)			
Quantity	Embedded Voice Extractor	Embedded Voice Client <sup>(1)</sup>	Embedded Voice Matcher
1-9	€ 9.00	€ 9.00	€ 11.00
10-19	€ 6.50	€ 6.50	€ 8.00
20-49	€ 5.80	€ 5.80	€ 7.10
50-99	€ 5.10	€ 5.10	€ 6.20
100-199	€ 4.60	€ 4.60	€ 5.60
200-499	€ 4.10	€ 4.10	€ 5.00
500-999	€ 3.60	€ 3.60	€ 4.40
1000-1999	€ 3.20	€ 3.20	€ 3.90
2000-3999	€ 2.90	€ 2.90	€ 3.50
4000-7999	€ 2.60	€ 2.60	€ 3.10
8000 and more	Please contact us for more information		
License management			
Volume license manager			€ 16.00

(1) These components are not available for VeriEye Standard SDK customers.

VeriSpeak products can be ordered:

- online, at [www.neurotechnology.com/cgi-bin/order.cgi](http://www.neurotechnology.com/cgi-bin/order.cgi)
- via a local Neurotechnology distributor; the list of distributors is available at [www.neurotechnology.com/distributors.html](http://www.neurotechnology.com/distributors.html)