Magazine
New Customer Management Insight Digital Publication
Call Center Magazine Archives
FREE Webcasts
Podcasts
Vendor Guide
Multimedia Content
Online Advice
eNewsletter Alerts
ICMI Forum
ICMI Jobs
Training
Consulting
Newsletters
Publications/Tools
Industry Research
QUEUETIPS
Conferences
ICMI Global Report
Call Center Knowledge Online
Join ICMI
Members Only
Vendor Guide
Industry Information
About ICMI
Download the ICMI Catalog

The Portal Pushers: Speech Vendors Poised To Grow The Voice Web

Voice access to websites started as a cute option on UM systems. Two years later, there's a voice mark-up language and a move to hosted voice sites that might actually grow into an industry.

By Ellen Muraskin

12/05/2000, 12:00 AM ET

Speaker Verification

V-commerce - the revenue-generating service that voice portals are hoping for - won't emerge until and unless consumers feel safe in using it. If e-commerce requests for credit card numbers made many e-shoppers pause, fear of wireless wiretapping may stop v-commerce cold. That's why verifying the identity of the unseen speaker/shopper is so important. Experts say that fraud can be largely prevented by three elements of proof of identity: something you know - a password; something you have - a credit account with a bank; and something you are. You are endowed with a unique voice print.

Matching entered voice to a stored voice print saves human resources and time: call center agents no longer have to ask callers for their mothers' maiden names, dogs' names, or other arcane information, or transfer them to human-staffed third-party verification services.

Automatic speaker verification is now used to access voicemail, unified messaging systems, calling card systems, cellular access and phone banking, in addition to credit-card purchase. It's even used to augment electronic surveillance by verifying the identity of individuals within restricted areas, or whose movements are restricted due to home detention or incarceration.

Buytel / ITT-SpeakerKey

For many years a division of ITT

UTILITIES
print this article
e-mail this article

RELATED LINKS
Why Is It Taking So Long For Speech Technology To Catch On?
Cultivate An On-Demand Workforce Through On-Demand Technology
Executive Interview: Intervoices Ken Goldberg
Nuance Releases Recognizer V9
Avaya Releases Customer Interaction Express
Spring Cleaning: Its Time to Clear the Cobwebs from Your Center
Nuance Unveils Voice Search
Microsoft to Acquire Tellme Networks

Aerospace/Communications, SpeakerKey (Fort Wayne, IN - 219-451-6321) now sells its products exclusively as an embedded part of Voicevault from Buytel Ltd. (Dublin, Ireland - 011-353-1-603-9500).

Datamonitor Reports On Speech Recognition

Nuance Acquires BeVocal

Voicevault is an ASP-like entrusted third-party service, available via telephone, the web or Internet, that authenticates customers by their voice and supplies support functions such as text-to-speech and database imaging. Voicevault services can handle existing phone and Internet interfaces, as well as new and legacy IVR systems. The service integrates automatically back to the applications of the service provider, who pays only for the particular services used per event - there's no cost for unused spare system capacity.

Buytel's Voicevault uses three voice biometric authentication products from ITT's SpeakerKey:

- **Phonekey** allows the user to place a call to a toll-free number, using any standard telephone handset (including cell phones), via the PSTN to an IVR program and a PhoneKey server, both located at an ITT/Buytel site. The IVR program is customized to conform to the service provider's specific application.

  Unlike a conventional IVR program that collects touch tones, this one collects speech samples and sends them to the PhoneKey server, which uses speech recognition and voice print comparison techniques to obtain the caller's account number and validate the caller's identity.

  Depending on the specific service offered, validated calls are then either forwarded to the service provider's site (via the phone network, for example); connected to a call center CSR; or data is taken by PhoneKey and transmitted back to the service provider (via Internet, for example), providing any kind of secure, remote data entry.

  If the caller's voice doesn't match the reference voice, the service provider can program PhoneKey to (1) ask the caller to repeat the voice input process, (2) transfer the call to an intervention operator, or (3) terminate the call.

- **Webkey** is a plug-in that enables a browser (running on a PC with a soundcard) to prompt a user to enter an account number and take a sample of the user's voice and authenticate the user before proceeding any further.

  The voice sample can be as simple as the account number itself. Alternatively, the verbal input can be repetition of random digit-pairs or password phrases, as prompted by the plug-in's GUI.

  The sample spoken phrases are converted to a special voice print format that's sent over the Internet to a server at an ITT/Buytel site.

The WebKey server then uses speech recognition and voice print comparison techniques to obtain the user's account number and validate the user's identity.

If the user's voice matches the reference voice print, a message is sent via the Internet to a Common Gateway Interface (CGI) at the service provider's website, allowing the user's browser to access protected pages.

WebKey can also provide digital recording and storage of calls for subsequent recall, if necessary (e.g., for stock transactions).

- **Netkey** services are similar to those of PhoneKey, but NetKey can be used with existing IVR platforms, thanks to the NetKey API, NKAPI.

  A small NKAPI software library installed on the legacy system provides everything needed to access Voicevault. Your existing IVR will operate normally, but will now prompt callers to say their account number and then record their response. Then, instead of calling a password procedure, the IVR makes a single procedure call to the SpeakerKey server, using NKAPI.

  For low-call volume applications, a server at ITT/Buytel is used. Higher-volume applications can justify purchase of a customer-site server.

  The Buytel/ITT-SpeakerKey server is a triply redundant system with an HP fiber channel RAID for voice print storage and multiple eight-way Xeon processors. The system is configured as multiple, completely independent, mirrored systems, located in different physical locations for maximum operational integrity. All sites include multiple telephone carriers and Internet gateways, utilizing multiple, redundant 100 Mbps copper, fiber-optic, and microwave services.

  Current system capacity (expandable) provides for 8 GB of main memory and storage for 64 million voice prints (at approximately 10 KB each). Each end-user verification process takes an average of 0.5 seconds.

Keyware Technologies

Speaker verification is just one aspect of Keyware Technologies' (Woburn, MA - 781-933-1311) Layered Biometric Verification (LBV) framework technology. The LBV Framework is an integrated solution for biometric verification starting from a one-password, one-trial verification process to the most sophisticated multi-trial, multi-password process, directed by a programmable API.

Keyware has put their voice verification algorithms into a network server that performs other biometric tests (for fingerprints or the face, for example), depending on clients' needs. The server checks previously stored voice prints against a voice password that's sent by an LBV-compliant client application.

The whole LBV Framework consists of: The LBV Server (a middleware application), Biometric Guardians (biometrics engine and data capture tools such as VoiceGuardian, FaceGuardian, and Finger Guardian), development tools, and application toolkits to speed up the development of applications in specific fast growing markets like telephony, smartcards, Internet, and physical access.

The LBV Framework is written in the form of internal and external software objects. The internal objects (COM objects grouped into an NT Service) run on the server, while the external objects run on the client. The external objects provide the interface between the LBV Client and the LBV Server and the LBV Framework can be called from any LBV Client. This client can be written in C, C++, or Visual Basic.

The LBV Framework can be used in a local or a distributed environment. Using DCOM, the client and the server can run on the same computer, on a network computer, or on a connected computer anywhere on the Internet. All client-server DCOM communication is encrypted, using the security built-in to Microsoft's DCOM protocol.

With Keyware's LBV technology, voice verification is a two step process. First, you enroll by repeating a "passphrase," at least two seconds long, three times. Good passphrase examples are "My voice is my password" or "Keyware Technologies," because they contain strong vowel sounds. Depending on the application, the user's voice profile is recorded into a database, a local PC, or onto a smart card.

Keyware sells a software development kit, VoiceGuardian, that allows developers to add ActiveX controls into apps that need voice security, including Web-based ones. Intended for applications that require more than just voice verification, the LBV server is sold separately and communicates with VoiceGuardian.

VoiceGuardian works independently of language using Lernout & Hauspie technology, can be text-dependent with user-selected passphrases, and is available in microphone and telephone versions. It has monitoring and usage reporting ability, enhanced front-end signal processing to ensure accurate recognition even with low-bandwidth signals, and small data requirements (3KB compressed for voice prints for standard storage, 1.8KB for low storage such as smartcards). It can check signal quality, giving users feedback on signal overload, bad signal-to-noise ratio, and insufficient volume.

Nuance

The Nuance Verifier 2.0 from Nuance (Menlo Park, CA - 650-847-0000) is an add-on to Nuance 7.0, Nuance's core speech recognition software.

Callers can either be enrolled and subsequently authenticated on the same utterance, or enrolled and authenticated on different utterances. For example, a caller can enroll saying their account number and then be authenticated based on a rotating question.

Applications can also offer seamless enrollment and authentication, where

both happen in the background. A caller may not even know that the application is performing a voice check while validating an entered ID number.

The Nuance Verifier 2.0 can take advantage of Nuance's distributed architecture, which supports both stand-alone and networked server configurations. Also supported are simultaneous load balancing of speech recognition, natural language understanding, voice authentication, and text-to-speech resources. Optimal use is made of every server CPU in the network, thus minimizing the hardware needed. Also, voice prints can be stored in multiple independent off-the-shelf databases if needed, to ensure scalability. These databases can be connected via a network, so that one centralized repository can serve multiple call centers.

Nuance claims that their Verifier has been measured to provide 99.9% security, with 94% call completion under real-life conditions. This means that there is less than a .01% chance that an imposter will be accepted into the system.

There is no limitation to the number of users that can be supported.

Nuance Verifier 2.0 supports North American English, UK English, Japanese, Latin American Spanish, Canadian French, European French, German, Swedish, Brazilian Portuguese, Cantonese, and Mandarin.

The verifier continually tweaks its voice print to optimize performance automatically. Nuance also packages up its voice authentication software into a SpeechObject, a reusable software component for the speaker verification developer's market.

T-Netix

T-Netix, Inc. (Englewood, CO - 303-790-9111), offers SpeakEZ voice print technology. Many speaker verification vendors employ a single template technology based on Dynamic Time Warping (DTW), Gaussian Mixture Models (GMM) or Hidden Markov Models, (HMM) which take measurements from the speaker only. SpeakEZ also uses a discriminant training-based pattern classifier called the Neural Tree Network (NTN).

NTN is a hierarchical classifier that has the properties of both decision trees and neural networks. During discriminant training, the NTN learns to contrast the voice of a given speaker with a select pool of "anti-speakers" (cohorts) with similar voice characteristics.

This approach yields higher accuracy (front-end analysis recognizes and normalizes conditions such as background noise, channel differences and microphone variances), needs less training data, and can take place in a relatively shorter period of time (100 to 200 milliseconds). There is no limit on the number of enrollees that can be recognized by the system.

Thus, SpeakEZ Voice Print technology can be deployed in the most difficult situations, such as cellular, public pay phone, and prison phone environments.

Speaker inputs for the system enrollment process include the user's

identity (i.e., PIN) and his or her password utterances. The output is the speaker's SpeakEZ Voice Print Speaker Model file (about 5KB for a 1-second utterance) and a Decision Threshold Score.

T-Netix technology has found its way into products of BioNetrix' Enterprise Biometric Management System; Envox's CT development environment; IBM's DirectTalk/6000 and DirectTalk/2 interactive voice response software; Lucent's Mobile Switching Center's Roamer Verification Reinstatement feature, to combat theft by clone or stolen phone in a roaming environment; Nortel's SecurPBX, Periphonics' OSCAR IVR platform, and Visionics' FaceIt.

T-Netix and Peak Network Communications have formed a joint venture to develop voice-based Internet security solutions. The new entity, Authentor Systems (Englewood, CO - 303-792-3726), develops speech rec products based on T-Netix's SpeakEZ Voice Print, VeriNet Web technologies, and Peak's Web server security software. Sentry's first product was VoiceKey, a software product that secures enterprise-level network communications and transactions with PC microphone-entered voice samples.

< < Previous Page | 1 | 2