Myths and Misunderstandings about Speaker Authentication

A white paper by
Dr. Judith A. Markowitz, president
J. Markowitz, Consultants

Introduction

Speaker authentication is a biometric technology that uses the acoustic sound patterns in a person's voice to ensure that that person is who they claim to be. Biometric-based security is different from other types of electronic security because biometrics provide direct authentication of the individual rather than authentication of something that person has or knows. Biometrics accomplish that by requiring that a person provide a biometric sample – a spoken utterance, a fingerprint, an iris scan, etc. – as proof that they are who they claim to be.

Because speaker authentication uses data from human speech it is often seen as less alien than some other biometric technologies. Because talking is an everyday skill speaker authentication is also often seen as less threatening than some other biometrics. There are still many questions and misconceptions about speaker authentication technology, how it works and whether it works.

This white paper provides a list of and answers to some of the most common questions and misunderstandings. Some of them apply to other biometrics and some deal specifically with speaker authentication. This is a living document so if you know of a myth or misconception about speaker authentication please let me know about it.

Myths and Misunderstandings about Speaker Authentication

Why do we need biometrics when PINs and passwords work just fine?

It's true that most of the electronic security in use today involves a user ID and a password or PIN. ID + password security is so ingrained in the culture of computer and network security that concerns about hacking, identity theft, and privacy have been met with programs requiring more arcane passwords with higher *entropy* that change more frequently and cannot be reused.

Aside from entropy, there is little proof that these those longer, more complex passwords actually heighten security. Instead, there is evidence that they actually reduce security. It is, for example, well known that users solve the problem of complex, ever-changing passwords by writing them down and storing lists of them nearby – sometimes tacked to their computer monitors. They persist even when instructed not to do so by security professionals because the new passwords are impossible to remember.

One of the most dramatic and compelling refutations supporting the effectiveness of ID + password security comes from the United States Department of Justice (DoJ). In 2007, the DoJ released *Network Attacks: Analysis of Department of Justice Prosecutions 1999 –*

2006. That reports describes the DoJ's analysis of ALL the cases of data network attacks that the agency prosecuted and publicly disclosed between March, 1999 and February, 2006. The cases involve everything including spreading malicious code such as viruses, Trojans, spyware and worms; theft of valuable data like intellectual property, credit card or other private financial information; and of other crimes, such as denial of services, unauthorized access, and financial theft. Each crime was examined and classified according to: type of attack; the methods used to carry it out; the attacker's relationship (if any) to penetrated organization; the location of attacker at time of attack; and the type, location and nature of equipment used in attack.

One of the key findings of this report pounds a stake into the heart of ID + password security:

> "organizations suffered the greatest financial loss and damage, more than $1.5 million per occurrence, when attackers used stolen IDs and passwords" (p.10)

The damage to individual companies ranged from thousands of dollars to millions of dollars. Despite the huge financial losses from these kinds of attacks, they are not sophisticated hacks. Most were perpetrated by individuals who were able to obtain valid user IDs and passwords.

Obtaining IDs and passwords is a flourishing criminal activity. Some of it is perpetrated using software (e.g., spyware, sniffers, and password generators) that is easily obtained on the Internet and elsewhere. Other thieves hack information repositories containing passwords and other sensitive information; still others wheedle the information out of users, call center agents, and others through the use of "social engineering."

So ID + password security is not effective, but it is costly in a far more pervasive fashion because it has spawned password reset. Password reset is the most-requested service provided by corporate and government help desks. Technical help desks spend, on average, forty to sixty percent of their time resetting passwords. This is costly for several reasons: the time of highly-paid technical professionals diverted to what is, from a technical perspective, very simple; those technical professionals have authentication and other security procedures added to their jobs as a part of ensuring that passwords are not given to criminals employing social engineering; and there is the cost of lost productivity of the individual whose password needs to be reset.

Ironically, the biggest market for speaker authentication today is to provide secure, automated password reset.

Why can't I just use speech recognition?

This question is a natural byproduct of the human ability to simultaneously recognition the speaker and what that person is saying. There is also confusion because of the unfortunate practice of using the term *voice recognition* as a synonym for *speech recognition.*

Unlike humans, speech recognition and speaker authentication are specialists.

> **Speech recognition.** Its goal is to answer the question "What are you saying?" This means that speech recognition focuses on words and phrases that are being spoken and is not at all interested in who is speaking. In fact, most speech recognition is *speaker independent*. That is, it contains as little speaker-specific information as possible so that it can handle a broad spectrum of speakers effectively. If you use speech recognition as part of a user authentication process you are essentially using a spoken version of traditional typed ID + password systems. No biometric component is involved.
>
> **Speaker authentication.** Its goal is to answer the question "Are you who you claim to be?" Speaker authentication is also called *speaker verification*) because its goal is to verify whether someone is telling the truth when they claim to be you. It achieves its goal by focusing on acoustic characteristics that distinguish your voice from the voices of other people. That is what makes speaker authentication a biometric technology.
>
> Some speaker-authentication technology can look for those aspects of your voice no matter what you say. That kind of speaker authentication is called *text independent*. Other speaker-authentication technology narrows the analysis by asking you to say something you've said to the system before, such as a voice password. That kind of speaker authentication is called *text dependent*. Some speaker-authentication technology requires you answer questions or to repeat a randomly-selected sequence of words or numbers. That kind of speaker authentication is called *text prompted*.

Speaker authentication operates well without speech recognition. In fact, most voice-authentication technology has been developed independently of speech recognition and is designed to be used by itself. Despite this, it isn't unusual to encounter both speaker authentication and speech recognition in the same application. The reason is that the differences between speech recognition and speaker authentication are complimentary. Speaker authentication adds security to a speech-recognition application without forcing the user to change modalities. In text-dependent and text-prompted applications speech recognition ensures that the person has actually said what the system asked them to say.

Why do I have to enroll?

Enrollment is generally the first encounter you have with a biometric system of any type. It occurs when you provide one or more samples of your speech to the system as part of a larger registration process.

Rather than being dispensable, a well-done enrollment is one of the most critical elements of a speaker-authentication deployment because

1. Enrollment provides the only samples of your voice that the system knows for sure come from you. These are the samples that are used to make the system's *reference model*. Your reference model is what is used to verify that someone claiming to be you is, in fact you.
   This means that the enrollment procedures must include the best user-authentication techniques that the organization deploying the speaker-authentication application can bring to bear. It may entail appearing at a pre-determined enrollment center; it may involve a special one-time password; it may include asking a number of questions only you should know; it may involve all of those things and more. After all, if the wrong person is enrolled under your identity then only that person can be authenticated by the system as you later on.

2. It is essential to use only good-quality samples to generate the reference model. This point falls under the garbage in-garbage out dictum. If the voice samples that you provide during enrollment are flawed in some way the quality of the authentication performed by the system later will be diminished.

Sometimes, it is possible to enroll someone when they are not aware of being enrolled. Generally, that is done for some applications of speaker identification which is asked to answer the question "Who is this speaker?" In contrast, speaker authentication requires that a user submit a claim of identity and tries to answer the question "Is this person who she/he claims to be?"

Isn't the "voice model" in a speaker-authentication just a recording?

No, it is very different from a recording. Speaker authentication doesn't need all of the information that makes up a full recording of your voice. It only needs to know about the characteristics of your voice that distinguishes it from other voices and makes it consistently your voice. This is why one of the first steps performed by a speaker-authentication system is *feature extraction*.

Feature extraction is the step during which critical features of your voice are located in the speech samples and coded. During enrollment, those features are used to create the *reference model* of your voice that is stored in the system's user database. Those features

are also used to create a model of the voice of the person claiming to be you. It is also those features in those two samples that are compared during the authentication process.

Once the critical features are removed from the voice sample the recording is no longer needed. An organization may wish to keep the recordings in case it needs to process the enrollment or verification interactions again. This may happen if, for example, the reference models in the database are corrupted or if the organization changes technology vendors.

Why is speaker authentication called a "behavioral biometric?"

The name *behavioral biometric* was given to speaker authentication, sign/signature dynamics, typing dynamics (also called *keyboard dynamics*) and other type of biometrics that require a person perform an act in order to create the biometric sample that a biometric system can capture and use. You can't analyze speech until a person begins to talk. This dynamic or volitional component distinguishes behavioral biometrics from *physical biometric*, such as iris recognition and face recognition because you can capture the image of many of those biometrics without having the person do anything. The distinction becomes a bit fuzzy for biometrics like vein, hand/finger geometry, or even live-scan fingerprint which are hard to capture in a purely passive fashion (That is, without asking the person to touch or approach a biometric sensor).

Most of the critical features that speaker-authentication systems extract from your voice provide information about the anatomy and physiology of your *vocal tract* rather than about your style of speech. As figure 1 indicates, your vocal tract starts you're your vocal cords and continues through your throat, mouth, and nose.

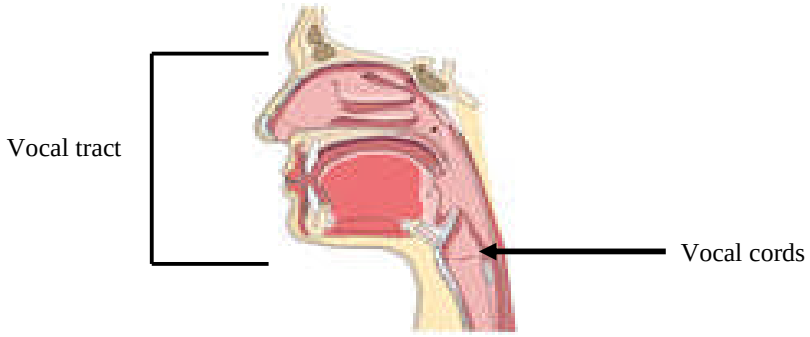

Figure 1 Human vocal tract

This emphasis on anatomy makes speaker authentication a physical *and* a behavioral biometric. It is also the reason why professional mimics have more difficulty fooling speaker-authentication systems than fooling people. Mimics simply do not have the proper vocal tract size and shape.

Another important point to remember that, in practice, virtually all commercial biometric authentication systems require the claimant to actively interact with the biometric sensor.

You must place your finger on the glass, look at the camera, or write on the tablet. This makes all biometrics behavioral as well as physical.

Awareness of and sensitivity to the behavioral component of biometrics is important for ergonomics and usability. For example, managing the behavioral element of submitting a biometric to ensure that the user can submit a good sample can prevent many unnecessary errors for all biometrics.

Why can't I expect biometrics to be 100 percent accurate?

The reason is that all biometric authentication is statistical.

When an authentication is performed by a biometric technology, such as speaker authentication a model created from the sample provided by the person claiming to be you and compared with your reference model (the model for your voice or other biometric stored in the system). That process is called *matching*.

It would be nice if matching could be performed on unchangeable samples but that isn't true for any biometric input. That's because biometric samples are provided in real-world environments that have variable temperature, humidity, noise, lighting and other factors; by humans who move their heads, place their fingers, and shape their tongues and mouths differently each time they interact with the biometric system; and by sensors that vary in quality, age, cleanliness, and position.

These and other factors result in variability in the input. Consequently, biometric authentication systems cannot look for exact matches. Instead, they generate a decision that expresses how similar two models are and use a similarity threshold as a cutoff point for accepting the claim of identity as valid.

Because biometric technologies are statistical they can make two kinds of errors:

1. *False match* (also called *false acceptance*) occurs when the system mistakenly accepts the identity claim of an impostor; and

2. *False non-match* (also called *false rejection*) occurs when the system mistakenly rejects your claim that you are you.

This means that no biometric is 100 percent accurate – even in the laboratory. This is as true for DNA as it is for speaker authentication.

It is also critical to remember that laboratory tests – even be reputable third parties, such as the US National Institute of Standards in Technology – cannot predict the performance any biometric will exhibit in your application. Laboratory tests simply demonstrate that a core technology works. Thus, it is unfortunate that in its report on the results of its 2006 test of face recognition products stated

> "In an experiment comparing human and algorithm [system] performance, the best-performing face recognition algorithms were more accurate than humans." (Phillips, J; Scruggs, W; Flynn, A; Bowyer, K; Schott, C; and Sharpe, M 2007 *FRVT 2006 and ICE 2006 Large-Scale Results* p. 1)

This level of performance cannot be guaranteed for real-world deployments where lighting, face position, and movement dramatically reduce the performance of those systems. Two recent examples come from Virginia Beach, Virginia and the German Federal Police (BKA) both of which terminated face-recognition programs. The three face recognition systems used in the BKA test. When conditions were optimal they achieved 60 percent accuracy.

The failure of face recognition in the Virginia Beach and BKA programs demonstrates that the environment in which a biometric system is deployed has a strong impact on accuracy. Those two programs were run in public areas. Other errors come from poor application design, poor ergonomics, and (as mentioned earlier) the variability that is inherent in real-world deployments.

Aren't other biometrics more accurate than speaker authentication?

It is hard to compare speaker authentication with other biometrics because voice and keystroke dynamics are the only biometrics technologies that were created – from the start – to work with standard input devices, such as telephones and keyboards. All other biometrics are designed to run on proprietary devices or third-party devices to which the technology has been ported. This means that any claims about performance are skewed in favor of biometric technologies that can control the characteristics and quality of their input devices.

Certainly, any technology will operate with greater accuracy when it is used with a device that is finely tuned to its needs. The National Physical Laboratory of the United Kingdom found this to be the case for voice as well as other biometrics. In fact, the National Physical Laboratory found that voice performs as well as or better than other biometrics when the "playing field" is made level by allowing the speaker-authentication system to exert as much control over the input device as the other biometrics. The results are displayed in table 1 (on the next page).

The brown line in table 1 describes the performance of the sole speaker-authentication product included in the test. It out-performed most of the other biometrics in terms of both false match and false non-match errors.

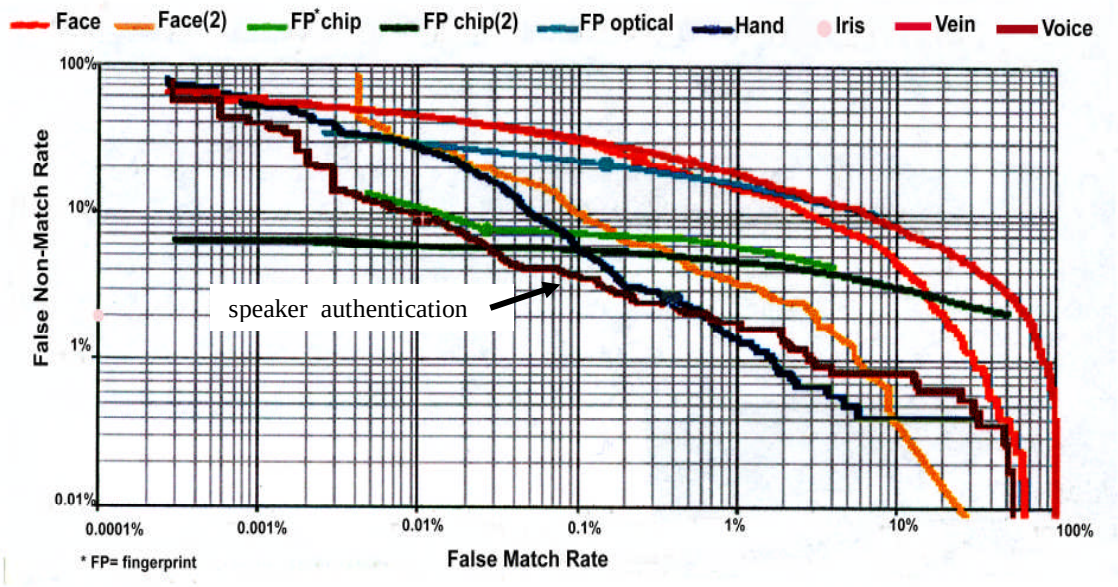

Table 1. Comparison of error rates for different biometrics) (*Biometric Product Testing - Final Report.* United Kingdom: Centre for Mathematics & Scientific Computing National Physical Laboratory. March 2001. P. 10)

Is speaker authentication ready for the real world?

Absolutely. In fact, speaker authentication has been operating in the real world for more than ten years. The earliest deployments were to monitor home-incarcerated offenders in community corrections and prison/jail inmates. Many of those deployments are still in use, including ones that employ an alcohol breathalyzer testing device with embedded speaker authentication. Like other community corrections applications, the use of the alcohol breathalyzer makes it possible to biometrically verify the location (and degree of intoxication) of the offender. Today, the use of speaker authentication for offender monitoring is a global phenomenon.

More recently, private industry has begun deploying speaker authentication. Among the most highly-visible of those deployments was the Dutch financial giant ABN AMRO's deployment of speaker authentication in its telephone-banking system and Bell Canada's nationwide deployment of speaker authentication on its automated customer-services lines.

Other private-sector deployments of speaker authentication include

> Aeroplan (Air Canada's rewards program), Allianz Dresdner, AT&T, Australian Health Management, Austar, Banco Bradesco of Brazil, Bank Leumi of Israel, Pershing, Morgan Stanley, Mitel, Progressive Corporation, US Bank, VeriSign, Wells Fargo Bank

Public-sector deployments include

> CentreLink (Australia), Commonwealth of Pennsylvania, Dubai Municipality, Illinois Department of Revenue, Maine Department of Corrections, Union Pacific Railroad, US Court of Appeals of New York, US Department of Homeland Security, US District Court of Florida.

Won't a tape recording fool speaker authentication?

The use of a tape recorder to try to fool a speaker-authentication system is called *spoofing*, *replay attack* or *tape attack*. A tape attack occurs when someone records your voice and then tries to fool the speaker-authentication system by playing that recording into the telephone.

Replay attacks against all biometrics are valid threats. Most speaker-authentication technology can spot tape attacks using standard recording equipment. High-quality digital equipment can fool voice-authentication systems just as fake fingers can fool fingerprint systems.

Although the threat posed by replay attacks is real, it is generally accompanied by two misconceptions that are used as reasons for not deploying speaker authentication. They are:

- It is easy to make a tape recording of your voice that can be used to attack a speaker-authentication system
- There is no way to detect or prevent replay attacks.

It isn't easy to create a reasonably good recording of the voice of an unaware and uncooperative victim. The recording must be done covertly but simply wearing a concealed microphone and standing next to the victim is unlikely to work -especially if that telephone is in a location where you would be out of place, such as the person's home. The same problem applies to placing a recording device in or near the victim's telephone. The best method is generally inserting the recording in the system at a point beyond the input device. The ability to do that is easier to use against VoIP than conventional public-switched telephone networks because criminals can use the same techniques that are used to attack other data networks.

There are many ways to detect and prevent replay attacks. Good voice-authentication systems look for the auditory signs of a replay attack. The most obvious signs are the telltale noise of conventional recorders, a match between the reference model (or another previously-captured model) that is too good, and the inability to break out of a limited set of utterances. Most commercial speech-authentication technology looks for matches that are too good to be true but detecting noise patterns is more challenging because tape

recorders vary in their noise characteristics and those patterns can, in some cases, be masked by other noise.

The best barriers are *liveness detection* and multi-factor authentication. The most potent liveness-detection measure is challenge response. The system may, for example, ask the claimant to repeat a random sequence of digits or words or even ask them to say something they've never said to the system before, such as "What's today's date?" A tape recorder could not reproduce a novel utterance like that.

A growing number of regulations and guidelines (e.g., FFIEC Guidance, HIPAA) are recommending or mandating multi-factor authentication, especially for higher-security operations. Multi-factor authentication can include other biometrics; knowledge verification (e.g., "What is your mother's maiden name?"); or even background operations, such as verification of the telephone or the location or behavior profiles.

Won't a professional mimic fool speaker authentication?

It is much easier for a professional mimic to fool you and me than it is to fool a speaker-authentication system. The reason is that mimics usually can imitate a speaker's style. Some are also able to modify their voices to approximate another person's pitch patterns but they can't change their anatomy or physiology. As mentioned earlier (see **Why is speaker authentication called a "behavioral biometric?"**), even though speaker authentication systems analyze some information that is related to a person's style of speech most of the information that is used defines the size and shape of the speaker's vocal tract. This also means that an amateur mimic who is a twins or close same-sex relative poses a far bigger threat to speaker authentication than a professional mimic.

Will it know me if I have a cold?

Most of the analysis done by speaker-authentication systems examines the size and share of your vocal tract which doesn't change much when you have a cold or the flu. Like speech recognition, however, speaker authentication relies on having rich acoustic data – lots of voiced sounds, such as *m. l, z, a, u* and *d*. When those sounds lose their power due to illness there is less information available for analysis. If the cold is severe and there is laryngitis the accuracy of speech recognition and speaker authentication will be affected. Minor congestion due to a cold should not have much affect on the operation of a speaker-authentication system, however.

Isn't the best password your own name?

Your name seems to be a logical choice as a spoken password. It is something you certainly won't forget and the level of familiarity is such that you will probably say your name in a fairly consistent fashion from one time to the next.

It turns out, however, that names are not especially good voiced passwords for speaker authentication. One reason is that we say our names so often that they are ideal candidates for replay attacks (see **You can fool a speaker- authentication system with a tape recorder**).

A more compelling reason is that they vary so much in their acoustic properties that they make the performance of speaker-authentication systems. Common names like Charlie Chan, Patty Page, Hal Holbrook, and George Bush are very short and filled with "voiceless" sounds, like 'ch,' 'p,' 't,' 'h,' and 'sh." Those sounds provide little "acoustic meat" for a speaker-authentication system to use in its analyses. This results in higher error rates, such as those shown in Table 2.

Test	**2006***		**2005****
	Persay commercial	Persay alpha	Nuance Verifier 3.5
Counting 1-9	1.05	0.77	0.91
Names	3.74	3/78	5.13

Table 2. Comparison of names and numbers
* University of Canberra 2006 Persay Technology Evaluation Results
** University of Canberra 2005 Speaker Verification Evaluation Report *Scientific evaluation of speaker verification technologies on behalf of Australian Government* Document No: SVE Test Report Version 2.0 Project ID: RFP-SV-026f

Table 2 contains the results of two sets of tests done by the University of Canberra comparing counting with personal names. The names were those of benefits recipients in Australia who participated in a speaker-authentication pilot for the Australian government. Table 2 shows that the error rates for all three of the products jumped when participants said their own names.

Can speaker Authentication be used to secure my garage door or my home?

Yes, it can. There are garage-door controllers, automobile locks, door locks, and other products and deployments that use speaker authentication for *physical access security*. In fact, in the 1990s, the US and Canadian immigration and customs services used speaker authentication for border control at an automated port of entry (POE) between the two countries. The system was deployed because families and businesses in the North American heartland around Scobey, Montana often straddle the border. After regular business hours the Scobey POE was unattended forcing local residents wanting to cross the border to travel up to 100 miles to the nearest open POE.

Both countries wanted to include a biometric in the automated POE they were constructing at Scobey but the POE is outdoors and weather conditions can be brutal with high winds and temperatures going lower than 60 F below zero. Under those conditions,

the only biometric that could be deployed was speaker authentication. An individual wishing to cross the border at Scobey would pull the telephone-like handset into the cab of their vehicle and interact with the speaker-authentication system. Once authenticated, they would replace the handset and cross the border.

Won't speaker authentication be used to invade my privacy?

The relationship between privacy and biometrics is complex and increasing convoluted. It is characterized by conflicting agendas (some of which are hidden or partially hidden), contradictions, and high emotions. Consequently, it is not sufficient to simply say that, by and large, speaker authentication does not invade your privacy.

In order to use speaker authentication you first must make a claim of identity and then you supply one or more voice samples that are compared with the reference model for your identity before allowing you to proceed. That is all it can do. Consequently, you speaker authentication can be seen as a technology that protects privacy and personal data because it erects an barrier to access by criminals.

Because of the high emotional component of the debate surrounding biometrics and privacy it is essential that any organization planning to deploy any kind of biometric, including speaker authentication, be mindful of privacy principles. The following set of privacy principles is called the *International Safe Harbor Privacy Principles*. It was drafted in 1999 by the United States Department of Commerce in collaboration with the European Union and must be signed by an American company before it is allowed to capture personal data on citizens of the European Union. The Safe Harbor principles are typical of many documents dealing with data privacy.

1. Notice: An organization must inform individuals about the purposes for which it collects information about them, how to contact the organization with any inquiries or complaints, the types of third parties to which it discloses the information, and the choices and means the organization offers individuals for limiting its use and disclosure. This notice must be provided in clear and conspicuous language when individuals are first asked to provide personal information to the organization or as soon thereafter as is practicable, but in any event before the organization uses such information for a purpose other than that for which it was originally collected or discloses it to a third party.

2. Choice: An organization must offer individuals the opportunity to choose (opt out) whether and how personal information they provide is used or disclosed to third parties (where such use is incompatible with the purpose for which it was originally collected or with any other purpose disclosed to the individual in a notice). They must be provided with clear and conspicuous, readily available, and affordable mechanisms to exercise this option. For sensitive information, such as medical and health information, information revealing racial or ethnic origin, political opinions, religious or philosophical beliefs, trade union membership or

information concerning the sex life of the individual they must be given affirmative or explicit (opt in) choice.(4)

3. Onward transfer: An organization may only disclose personal information to third parties consistent with the principles of notice and choice. Where an organization has not provided choice because a use is compatible with the purpose for which the data was originally collected or which was disclosed in a notice and the organization wishes to transfer the data to a third party, it may do so if it first either ascertains that the third party subscribes to the safe harbor principles or enters into a written agreement with such third party requiring that the third party provide at least the same level of privacy protection as is required by the relevant safe harbor principles.(5)

4. Security: Organizations creating, maintaining, using or disseminating personal information must take reasonable measures to assure its reliability for its intended use and reasonable precautions to protect it from loss, misuse and unauthorized access, disclosure, alteration and destruction.

5. Data integrity: Consistent with these principles, an organization may only process personal information relevant to the purposes for which it has been gathered. To the extent necessary for those purposes, an organization should take reasonable steps to ensure that data is accurate, complete, and current.

6. Access: Individuals must have [reasonable] access to personal information about them that an organization holds and be able to correct or amend that information where it is inaccurate. [Reasonableness of access depends on the nature and sensitivity of the information collected, its intended uses, and the expense and difficulty of providing the individual with access to the information.](6)

7. Enforcement: Effective privacy protection must include mechanisms for assuring compliance with the safe harbor principles, recourse for individuals to whom the data relate affected by non-compliance with the principles, and consequences for the organization when the principles are not followed. At a minimum, such mechanisms must include (a) readily available and affordable independent recourse mechanisms by which an individual's complaints and disputes can be investigated and resolved and damages awarded where the applicable law or private sector initiatives so provide; (b) follow up procedures for verifying that the attestations and assertions businesses make about their privacy practices are true and that privacy practices have been implemented as presented; and (c) obligations to remedy problems arising out of failure to comply with these principles by organizations announcing their adherence to them and consequences for such organizations. Sanctions must be sufficiently rigorous to ensure compliance by organizations.

Aren’t biometrics the Biblical “mark of the beast”?

This is another highly-emotional topic. It is the belief that biometrics are the “mark of the beast” that is discussed in the book of Revelations in the New Testament. It refers to a mark that is placed on your forehead or right hand by the charismatic leader of an international government and serves to control all individuals bearing the mark.

> "He also forced everyone, small and great, rich and poor, free and slave, to receive a mark on his right hand or on his forehead, so that no one could buy or sell unless he had the mark, which is the name of the beast or the number of his name. This calls for wisdom. If anyone has insight, let him calculate the number of the beast, for it is man's number. His number is 666." (Revelation 13:16-18)

Acceptance of the mark of the beast is a violation of God’s wishes

> "If anyone worships the beast and his image and receives his mark on the forehead or on the hand, he, too, will drink of the wine of God's fury, which has been poured full strength into the cup of his wrath." (Revelation 14:9)

Some deeply religious Christians consider some biometrics the mark of the beast. That belief generally does not extend to speaker authentication, but it may, and those religious beliefs should be considered when a biometric application is being designed.

Will my users accept biometric technology?

In the 1990s biometrics were seen as belonging in the realm of science fiction and any mention of using biometrics sent shivers down the backs of many consumers. It was a time when the use of facial recognition to spot pickpockets and other criminals in a crowd of Superbowl fans produced a public outcry.

Today, biometrics are not only becoming part of our lives, consumers support their use to protect against identity theft, fraud, and terrorism. In 2006, the Ponemon Institute, a privacy research institution, and UNISYS surveyed 16,683 adults in North America, Europe, Asia-Pacific, and Latin America

They found that, worldwide, consumers are concerned about identity theft and

- 69% support use of biometric technologies by a trusted organization
- 84% acceptance rate for voice (second highest)
- 32% listed voice as most favored biometric (highest)
- 66% favored biometrics as the ideal method to combat fraud and identity theft
- Convenience is the No. 1 reason (82%) for support of biometrics

They also found that

> "The most preferred biometric methods are voice recognition and fingerprints, and the least preferred method is a scan of the iris or eye." (*Global Study on the Public's Perceptions about Identity Management* p. 5)

Even though these results are typical of surveys that ask consumers about biometrics any deployment of speaker authentication needs to include a strong communication component that shows that your organization is sensitive to privacy and to your customers' or employees' issues.

About the Author

Dr. Judith Markowitz is recognized internationally as the leading independent industry analyst in voice-based biometrics and as one of the top analysts in speech processing. In 2003 she was voted one of the top ten leaders in the speech-processing industry and, in 2006, she was elevated to Senior Member status in the IEEE.

For over twenty years, Judith has provided strategic and technical consulting to organizations ranging from two-person startups to multinationals with 100,000 employees. She has also been actively involved in the development of standards in both biometrics and speech processing , including *ANSI X9.84 Management and Security of Biometrics in Financial Services*, *Speaker Verification API (SVAPI)*, *BioAPI*, and *ANSI/HFES 200 Human-Factors Standards for Speech Recognition*. She currently co-chairs the VoiceXML Speaker Biometrics Committee and is an invited expert in the W3C Speaker Identification/Verification Working Group. Her current work on standards include being the VoiceXML Forum liaison to ANSI/INCITS/M1 and co-editing a VoiceXML Forum-M1 project to develop a standardized data exchange file format for speaker verification and identification.

Judith publishes and speaks extensively in speech processing and biometrics and she's often quoted in speech industry and business publications. She's published over fifty articles in speech-processing, biometrics, and computer trade journals, such as "Voice Biometrics: Speaker Recognition in the Real World" in *Communications of the ACM*, and she's surveyed the voice-biometrics industry for both *Biometric Technology Today* and *Speech Technology Magazine*. Judith's publications also include the internationally-acclaimed book *Using Speech Recognition*, white papers, and the *Voice ID* series of reports on speaker-verification. She also writes a feature column for *Speech Technology Magazine.* She recently presented keynote speeches at Australia's National Centre for Biometric Studies Conference on Voice Authentication for Identity Management and the AVIOS Israel annual conference.

Judith has a doctoral degree in linguistics from Northwestern University, a Master's degree in computer science (specialization in AI) from DePaul University, and a certificate in strategic management from the American Management Association.

About J. Markowitz, Consultants
J. Markowitz, Consultants (JMC) is a sole proprietorship formed in 1990 for the purpose of promoting development of creative and profitable speech recognition, voice-biometrics, and knowledge-based businesses. JMC is now one of the leading independent industry analysts in speech-processing and biometrics.

JMC publishes three monthly monitoring reports on speaker authentication and other biometrics:

Program A: competitive landscape and standards
Program B: legislative, regulatory, and judicial activities
Program C: government activities

Contact us at contact@jmarkowtiz.com or visit www.jmarkowitz.com to find out more and to get a trial three-month subscription.

In our consulting work, JMC provides a framework for thinking through problems and for capitalizing on emerging opportunities. We bring you an independent perspective supported by extensive experience whether you need assistance in defining your requirements, selecting a speech-technology vendor, determining whether to invest or acquire a company, or even simply understanding what speech technologies can and cannot do.

We are based in Chicago and can be reached at

Web: www.jmarkowitz.com
Telephone: +1-773-769-9243
Fax: +1-773-769-9253
Email: contact@jmarkowitz.com

We look forward to hearing from you.