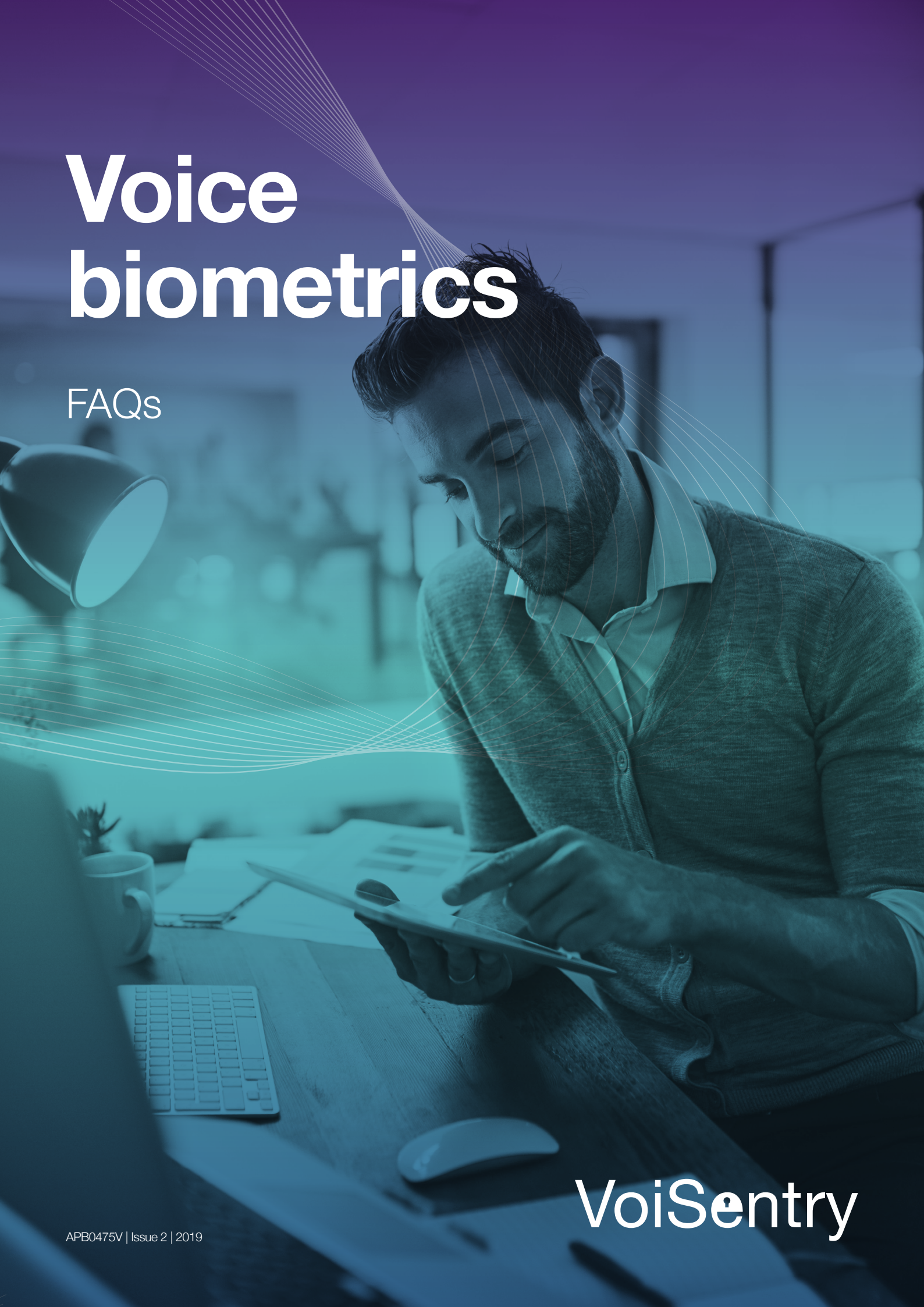
Voice biometrics

FAQs



VoiSentry

APB0475V | Issue 2 | 2019

What is voice biometrics?

Biometrics is the science of measuring and analysing human biological factors. Therefore, voice biometrics involves measuring and analysing a person's voice. The technology involved works on the principle that an individual's voice is unique, in the same way that a fingerprint and other biometric characteristics are unique.

The uniqueness of a person's voice is their physical characteristics and their speech habits. Due to that uniqueness, the identity of a speaker can be established or verified through an analysis of their voice. When a user registers in a system by providing samples of their voice, personal voice patterns are extracted from the audio and a unique reference model is created. The model or template is called a voiceprint, which is analogous to a fingerprint.

To identify or verify that a person is who they claim to be, voice biometrics technology uses algorithms to analyse their speech and compare that to the previously created model. If there is a match, the voice biometrics system will confirm that the speaker is the person registered against the voiceprint. Using a person's voice in such a way has become an accepted and established practice.

What are the main uses of voice biometrics?

Voice biometrics is used in two main ways; for the purposes of authorisation and identification.

The important use cases are:

a) to validate the identity of a person making an identity claim;

b) to identify who an individual is.

The first case is proof of identity, and is synonymous with providing a PIN or password to authorise access to a service. Instead of providing their password, a user of the system gives a sample of their voice. That sample is then analysed and compared with a voiceprint registered to the genuine user. This use case is commonly referred to as speaker verification. We can say also that the user's identity has been authenticated.

The other use case is to determine the identity of a person when e.g., they are claiming to be someone else or attempting to remain anonymous. By analysing a sample of speech and comparing that with all voiceprints in a database or hotlist, we can identify the individual. That's why this use case is referred to as speaker identification. It is particularly useful for detecting and identifying known fraudsters, or for recognising repeat malicious callers

Is voice better than other biometrics factors?

Voice is comparable to other biometrics in many ways. However, voice does have some advantages, not least because the user doesn't need a scanner, such as for iris and fingerprint recognition.

Voice is extremely easy to use, and, because of that, it has a higher level of user acceptance than many other biometric identity verification methods. In terms of accuracy, voice is broadly equivalent to other methods, and it is no less secure than fingerprints, retina, or facial recognition. Essentially, voice is both convenient and reliable, not having to be concerned with residues or poor lighting.

A primary advantage of voice is that it is the only biometric technology that can be used remotely over the telephone. That means it is particularly suitable for authenticating callers to a contact centre or an inbound, IVR-driven self-service platform. In a similar way, it can be used to verify the identity of people called from an outbound contact centre solution.

How does voice biometrics compare to traditional forms of communication?

Current, non-biometric methods involve shared secret knowledge and physical tokens. Secret knowledge takes the form of a PIN or password, or the answer to a security question i.e., it's something you know. Examples of physical tokens i.e., something you have, include keys, ID cards, security fobs, drivers' licences, and passports.

Unfortunately, the traditional methods are vulnerable to social engineering and theft. Tokens are routinely counterfeited and stolen, and passwords are routinely forgotten, left in plain sight, and stolen. Moreover, tokens can't guarantee the positive identification of a person.

In contrast, biometrics is less open to being copied, hacked, shared, or stolen. And aside from jokes about losing your voice, as it involves something that relates to who you are, it can't be lost. Furthermore, in terms of the inherent security of voice biometrics, a voiceprint is a derived code, it's not a recording, it can't be reverse engineered to reproduce speech, and if it were to be accessed by a hacker, the data would appear as a meaningless string of numbers that is functionally useless.

Using a biometric in combination with other methods equates to strong, multi-factor authentication. Having e.g., a mobile phone, knowing your PIN, and verifying your identity via voice biometrics is a secure method of reducing the vulnerability of systems and services to unauthorised access.

How accurate is voice biometrics?

First of all, no biometric system is 100% fool proof. Industry reports and studies indicate that success rates above 90% should be the minimum acceptable, where success means that a person is able to authenticate their rightfully claimed identity.

Regardless of competitive claims, a clue to the nature of accuracy lies in the existence of three important errors that reflect the performance of a system. Those measures are the equal error rate (EER), the false acceptance rate (FAR), and the false rejection rate (FRR). Allowing an impostor to get into the system is an error of false acceptance. Denying a genuine user access to the system is a false rejection error.

Industry reports indicate that optimal, text-dependent voice biometric engines can achieve a FAR of below 1%, with a corresponding FRR of less than 3%. However, it is difficult to compare systems based on published 'accuracy' figures. That's partly because there is no industry standard dataset against which to measure performance, and partly because threshold settings and similarity ratings are vendor and implementation specific.

Notably, the point at which the FAR and FRR curves intersect is the EER. Thus, EER is the commonly accepted metric used to compare the separability of systems i.e., their effectiveness at differentiating between genuine users and impostors. That's because unlike FAR and FRR, EER is independent of the threshold setting.

Notwithstanding all that, the sensible thing to do in relation to voice biometrics is to run trials in your target environment, based on real-world users.

If voice biometrics isn't 100% accurate, why should it be used?

In spite of its imperfection, voice biometrics is a valuable technology in several ways. All businesses know about risk assessment, particularly in relation to IT and data security. If you conduct a risk assessment, the output will be preventative actions to mitigate the risk(s). Managing a risk means prioritising actions appropriate to the risk level.

Voice biometrics confers benefits beyond fraud detection and the mitigation of risks. However, in relation to mitigating fraud, voice biometrics offers a valuable resource in that:

a) its presence deters fraudsters;

b) its use makes the fraudster's task very much harder; and

c) the results of using it mean the business will lose less to fraud.

Consequently, it must be obvious that voice biometrics can be very effective in helping to manage fraud risk.

The vulnerability of PINs, passwords and security questions has led to some very public data breaches at well-known organisations, putting customers at risk. Voice biometrics can reduce the security risk and mitigate fraud, whilst offering a more convenient user experience.

However, the best practice is to implement voice biometrics in combination with other methods, leading to a strong, multi-factor authentication solution that will reduce the vulnerability of systems and services to unauthorised access.

What other benefits will I get from implementing voice biometrics?

The business case for voice biometrics centres around a four-fold benefit, namely:

1. Fraud and security risks are mitigated;
2. The cost of authentication is reduced;
3. The customer experience is improved; and
4. Call takers' morale and motivation is heightened.

In addition, there is a clear return on investment (ROI). An ROI is an accounting measure, but there can be other, 'non-bean-counter' benefits, which include the security outcomes, enrolment take-up, and customer satisfaction.

The benefits of increased security mean that:

1. you will lose less to fraud;
2. you will save by not having to pursue fraudsters through the legal system; and
3. you will save through having to pay out less in compensation and reimbursement.

The cost reduction benefit comes from replacing manual authentication with an automated system.

The more the process is automated, the greater the saving. A voice biometric system can shave two-thirds or more off the cost of identifying and verifying callers. In UK contact centres, for example, the average time to authenticate a call via an agent is around 30 seconds. Contrast that with an automated, voice biometric system that achieves the same result in 10 seconds or less and you can see the potential for savings.

Automating the authentication process undoubtedly makes it easier and more convenient for customers. In addition, removing the tedium of having to ask the same security questions, day after day, is bound to have a positive effect on call takers' morale. Customers will surely recognise the security benefits to them and prefer using technology to having to remember the name of their first pet (school/car/etc).

How those benefits are measured will depend on the business. However, in most scenarios a subscription licensing model will give you an ROI in less than a year, year after year after year.

What about languages?

Fortunately, language isn't a big issue, because voice biometrics works on the sounds that people make, rather than on what they say.

Some vendors will need to train their voice biometric engine to get the best results for each language, because different languages use different sets of phonemes, or produce a different engine for each language. However, the best systems will operate independent of language, because they cater for a wide range of language sounds.

That's not to say that a system can't be fine-tuned, but training a system in such a way can involve more than just language. It's feasible for a system to be fine-tuned for a fixed, text-dependent passphrase and for the speaker domain (e.g., environment, networks, and devices), but the most beneficial effects can be gained by applying best practices during set-up and implementation.

Where language is a factor, is where speech recognition is used in tandem with speaker verification (speaker recognition) to validate what is said in addition to who said what. However, that is a separate issue, which has no bearing on the performance of the voice biometric engine.

Follow VoiSentry: Twitter | Blog | LinkedIn

+44 (0) 1908 27 38 38 | www.voisentry.com

VoiSentry