customer care solutions
from Nuance

Solutions enterprise: FAQ

Biométrie Vocale

A propos de ce Document

Les solutions de biométrie vocale Nuance sont conçues pour répondre aux défis en matière de sécurité des institutions financières, des fournisseurs de services de télécommunications, des acteurs de la santé, des entreprises et du secteur public.

Les solutions Nuance, au travers notamment de l'identification et de l'authentification des utilisateurs de centres de contact, terminaux mobiles et applications Web, satisfont aux exigences de sécurité les plus strictes de respect des réglementations, de sécurisation des transactions mobiles, et de prévention de la fraude et de l'usurpation d'identité.

Ce document aborde les questions les plus fréquentes au sujet de la biométrie vocale .

FAQ

Qu'est-ce que l'authentification biométrique vocale, la vérification du locuteur, ou la signature vocale ?

Qu'est-ce que l'authentification biométrique vocale, la vérification du locuteur, ou la signature vocale ?

C'est la vérification automatisée de l'identité d'une personne, basée sur les caractéristiques uniques de la voix.

L'empreinte vocale est comparable à l'empreinte digitale à bien des égards. Une empreinte vocale combine la mesure des caractéristiques comportementales de la façon dont la personne parle, ainsi que celle des caractéristiques physiques de l'organe vocal de la personne.

Au cours de l'authentification de la voix, la voix de l'utilisateur est comparée à l'empreinte vocale stockée pour vérifier l'identité déclinée. L'utilisation de la biométrie vocale offre de nombreux avantages, parmi lesquels :

- elle est plus fiable qu'une carte magnétique ou un mot de passe qui peuvent être dérobés,
- elle s'intègre de façon transparente dans une conversation téléphonique
- elle ne nécessite pas de disposer de coûteux dispositifs de lecture ou de scanners, juste d'un micro ou d'un téléphone. Elle est utilisable à distance et sans contact.

La biométrie vocale permet-elle d'identifier une personne ?

L'empreinte permet en effet d'identifier une personne dans un groupe. Elle permet par exemple d'identifier l'une des personnes d'un foyer ou d'une petite entreprise, qui partagent un même numéro de téléphone ou un même numéro de compte. Chaque personne du groupe doit s'enrôler séparément, avec les mêmes identifiants. Ensuite, lors de la vérification en cours d'appel, Vocal Password est capable de vérifier que l'appelant fait partie du groupe, mais aussi d'identifier quel est le membre du groupe qui appelle.

En revanche, la biométrie vocale ne permet pas de rechercher l'identité d'une personne en temps réel parmi un très grand nombre (par exemple des millions d'empreintes), mais elle permet de réaliser l'authentification d'une personne dont l'identité est déclarée parmi une population de plusieurs millions d'empreintes enregistrées.

L'empreinte vocale ne remplace donc pas l'identifiant d'une personne, mais peut en revanche se substituer au mot de passe, ou venir le compléter pour une authentification multi-facteur, selon le besoin de sécurité.

Comment se passe l'enregistrement initial de l'empreinte vocale ?

Pour être capable de reconnaître l'identité de quelqu'un sur la base de sa voix, le système de

vérification vocale biométrique doit au préalable enregistrer la voix du locuteur. C'est ce qu'on appelle l'enrôlement. L'enrôlement se déroule lors d'un premier appel où le locuteur doit répéter sa phrase clé plusieurs fois (entre 3 et 5 fois environ). À partir de cet échantillon le système crée l'empreinte vocale proprement dite.

Il est crucial que l'identité de l'appelant lors de l'enrôlement soit vérifiée de façon extrêmement fiable, pour éviter qu'un imposteur puisse s'enrôler en usurpant l'identité du bénéficiaire. Dès la phase d'enrôlement terminée, le système est prêt à vérifier l'identité de l'appelant sur la base de sa voix et de l'empreinte vocale collectée.

Quelle durée de parole est requise pour l'enrôlement ?

Typiquement, une phrase clé de quelques secondes prononcée trois fois suffit dans la phase d'enrôlement. L'enrôlement est assuré par un serveur vocal ou une application mobile. La phrase ne doit pas être trop courte ; elle doit comporter au minimum entre une dizaine et une quinzaine de mots.

Avec la technologie FreeSpeech, lorsque l'enrôlement est effectué au cours d'une conversation avec un agent, on utilise généralement deux échantillons de 30 secondes, collectés lors de deux appels différents.

Une phrase clé doit-elle obligatoirement être utilisée, ou peut-on reconnaître quelqu'un sur la base de paroles différentes ou inconnues du système ?

Il est en effet possible d'utiliser des techniques de vérification indépendantes du texte. Celles-ci nécessitent en général des phases d'enrôlement plus longues et plus complexes.

Vérification indépendante du texte : elle est réalisée en utilisant seulement la voix du locuteur, indépendamment des paroles prononcées.

Vérification par mots-clés : la vérification est réalisée en demandant au locuteur d'énoncer des mots spécifiques constitués d'un sous-ensemble aléatoire composé de groupes de mots ou de combinaisons de mots utilisés lors de l'enrôlement.

Quelle durée de parole est nécessaire pour la vérification ?

Dans le cas de l'utilisation d'une phrase clé unique, une seule prononciation de la phrase clé est en général nécessaire.

Comment s'assurer que lors de l'enregistrement initial de l'empreinte, le système n'a pas affaire à un imposteur ?

On utilise pour cela différentes méthodes, telles que l'utilisation d'un code unique provisoire, un questionnaire personnel détaillé, ou une combinaison de méthodes classiques d'identification et d'authentification existantes. Pour plus de sécurité cette phase peut être assistée par un agent.

Comment se passe l'authentification dans le cas d'un numéro de compte utilisé par plusieurs personnes, par exemple pour un compte bancaire joint entre deux époux ?

Voir réponse à la question : "La biométrie vocale permet-elle d'identifier une personne ?"

Quel niveau de sécurité la biométrie vocale peut-elle me garantir ?

Très bon, et comparable à celui offert par d'autres technologies biométriques.

La performance réelle est liée à plusieurs éléments: le type d'application, sa conception (en particulier l'ergonomie des procédures d'enrôlement et de vérification), le réglage de paramètres, et l'utilisation de fonctions avancées comme celles offertes par notre solution Vocal Password.

Lorsque tous ces facteurs sont optimisés, une application Vocal Password peut automatiser l'authentification d'un grand nombre d'appels à des niveaux de sécurité très élevés.

Comment mesure-t-on ce niveau de sécurité ?

La performance d'authentification de la voix, comme pour toute solution de biométrie, est mesurée en fonction de deux taux d'erreurs distincts :

- Taux de fausse acceptation (ou FA): pourcentage des imposteurs qui parviennent à s'introduire dans le système;
- Taux de faux rejet (ou FR): pourcentage d'utilisateurs valides qui sont rejetés;

Ces taux d'erreur sont mesurés sur deux populations distinctes. Le taux de faux rejet est mesuré sur toute la population d'utilisateurs valides. Le taux de fausse acceptation est mesuré sur la fraction des appelants qui sont des imposteurs et qui cherche activement à forcer le compte de quelqu'un en particulier.

En pratique, un taux de FA de 1% ne signifie donc pas que 1% de la population entière serait accepté à tort, mais seulement que 1% des imposteurs cherchant activement à obtenir l'accès au système seraient autorisés.

En conditions de laboratoire avec des enregistrements audio de bonne qualité - nous obtenons moins de 1% du taux d'erreur équivalant (EER). En conditions réelles, avec des appels sur divers canaux et des conditions habituelles de niveaux de bruit, nous parvenons à un taux d'erreur aux alentours de 3%.

Est-ce suffisamment fiable ?

Rappelons que pour les applications sensibles, la biométrie vocale n'est jamais utilisée seule. En conjonction avec un code PIN par exemple, la biométrie vocale permet de réduire le risque d'usurpation d'identité à des niveaux extrêmement faibles.

La technologie est aussi utilisée pour sa simplicité et rapidité d'utilisation : en Australie, les abonnés de la télé par satellite peuvent commander de la vidéo à la demande pour de petites

sommes. Ils sont authentifiés par la solution de Nuance. Les abonnés sont identifiés par leurs numéros de téléphone, puis authentifiés par la voix, sans l'usage d'un mot de passe.

Faut-il toujours privilégier un bas taux de fausse acceptation plutôt qu'un bas taux de faux rejets ?

En pratique, il y a un compromis entre ces deux taux d'erreur à déterminer. Selon les besoins de l'application - Sécurité ou Commodité - Nuance Vocal Password peut être calibré pour s'adapter à chacun de ces besoins.

Par exemple, la plateforme Vocal Password peut être calibrée pour des applications de haute sécurité (moins de FA, plus de FR), pour plus de commodité (plus de FA et moins de FR) ou bien équilibrée, avec des taux de FA et FR similaires, aussi appelé taux d'erreur équivalant (Equal Error Rate - EER).

Nuance offre un outil, Evaluation Studio dédié à la mesure facile ces taux.

L'outil offre une méthodologie qui assiste l'utilisateur par étapes permettant de procéder à l'analyse détaillée des performances du produit et d'évaluer ses possibilités via différent scénarios. Il permet en particulier de simuler l'impact de la modification des seuils sur les taux de faux acceptations et faux rejets.

En résumé, Evaluation Studio permet sur la base de données réelles de choisir la configuration la plus adaptée à l'atteinte des objectifs de simplicité d'utilisation et des exigences de sécurité.

Quel est le niveau de performance relatif de la biométrie vocale par rapport à d'autres technologies de Biométrie ?

Bon et même excellent dans certains cas. Toute solution de biométrie doit faire face à un type de bruit dans son application. Les empreintes digitales par exemple ne sont pas très robustes. Non seulement la technologie est sensible à la saleté ou à des substances grasses sur le doigt ou sur le scanner, mais il est démontré que les empreintes digitales peuvent subir des dommages lors de l'utilisation de produits ménagers par exemple, mais aussi dans certaines activités agricoles qui peuvent complètement effacer les empreintes. De plus, certaines personnes ont une peau naturellement sèche, ce qui empêche tout prélèvement d'empreintes nettes. La reconnaissance faciale est aussi très sujette à différents facteurs comme la pilosité, ou bien les différences de conditions de luminosité lors de l'enrôlement et du contrôle.

La voix, quant à elle, est sensible au bruit ambiant et à la variabilité introduite par les différents canaux de communication et terminaux utilisés, bien que des progrès aient été réalisés dans ce domaine (cf question / réponse suivante).

Quels sont les avantages de la biométrie vocale par rapport à d'autres technologies de Biométrie communes ?

L'authentification vocale offre également des avantages importants sur les autres types de contrôles biométriques en termes de facilité de déploiement et de convivialité.

Facilité de déploiement: pas besoin de matériel de capture coûteux ou complexe. A grande échelle, les applications de consommation grand public comme la banque, le courtage, ou des services de télécommunications, offrent une méthode d'authentification qui peut être utilisée par tout le monde depuis la maison, le bureau ou en voiture, ce qui constitue aujourd'hui une exigence essentielle. Comme l'authentification vocale se fait à partir d'un téléphone, d'un mobile ou d'une tablette, c'est la seule technologie de biométrie qui peut être mise en œuvre simplement et rapidement pour tous les clients de l'entreprise. Par ailleurs, les utilisateurs peuvent s'enrôler dans le système et être vérifiés pratiquement sans assistance. La plupart des autres solutions biométriques nécessite des opérateurs bien formés et des conditions contrôlées pour la collecte correcte des empreintes biométriques.

Convivialité: le contrôle de la voix est aussi beaucoup moins intrusif pour l'utilisateur final que le contrôle de l'iris, du doigt, ou l'utilisation d'un scanner facial. Les gens se sentent mal à l'aise dès lors que leurs yeux ou leur visage sont scannés. Les empreintes digitales sont légèrement moins intrusives, mais nécessitent tout de même un contact physique avec un périphérique. La voix, au contraire, peut être parfaitement intégrée dans un dispositif qui peut être perçu de manière tout à fait naturelle à l'utilisateur.

Le type de téléphone que j'utilise a-t-il une influence sur le résultat ? Si je me suis enregistré sur un fixe et que j'appelle depuis un mobile, est-ce que j'ai moins de chances d'être reconnu ?

Oui, en effet, ceci peut avoir un impact. Ce sujet, que nous appelons "Cross Channel" est en effet un défi pour toute solution de biométrie vocale. Nuance dispose de son propre processus d'optimisation et de calibration qui nous permet d'ajuster le système pour avoir des résultats homogènes dans ces conditions cross-canal.

Récemment, nous avons lancé un nouveau moteur algorithmique qui repose sur un mécanisme de compensation par type de canal. Cet algorithme "apprend" les altérations caractéristiques de la voix propres à chaque canal et à partir de nombreux appels est capable d'annuler l'impact du canal sur la qualité du son.

Que se passe-t-il si j'ai attrapé froid ? Le système va-t-il me rejeter ?

Ceci, tout comme l'âge ou des altérations liées à un état émotionnel ou une pathologie, peut influer sur le résultat individuel. Mais même dans ces conditions le moteur est toujours capable de faire la différence entre vous et un imitateur ou un fraudeur.

Si quelqu'un arrive à se procurer un enregistrement de ma voix, peut-elle l'utiliser pour s'enregistrer à ma place ?

Nous avons un mécanisme de détection anti-reproduction (‘playback detection’) inhérent qui compare chaque segment audio à l'empreinte des n dernières vérifications, ce qui permet de déjouer facilement de telles tentatives utilisant des enregistrements.

Pour les applications nécessitant le plus haut degré de sécurité, Nuance propose également des techniques de ‘détection de présence’ ('liveness detection').

Est-ce qu'un imitateur talentueux comme Nicolas Canteloup pourrait se substituer à une personne authentique ?

Non, l'empreinte reflète des caractéristiques biométriques impossibles à reproduire, et même si l'oreille humaine a des difficultés pour les distinguer, le moteur fait nettement la différence.

Et une voix de synthèse, générée par ordinateur ?

Notre moteur a déjà été testé contre un moteur de synthèse par le passé, lors d'un benchmark indépendant, où il a été mis en évidence que notre système distingue très facilement la différence entre la voix originale et la voix synthétisée.

Est-ce que ça fonctionne dans un environnement bruyant, ou si des personnes parlent à côté de moi ?

Il est recommandé pour la prise d'empreinte de parler dans un endroit calme et silencieux. C'est particulièrement vrai pour l'enrôlement. Le moteur est toutefois suffisamment robuste pour supporter des conditions sonores que l'on rencontre fréquemment lors des appels depuis les mobiles. Mais on constate que le taux de faux négatifs (appelants authentiques rejetés) augmente dans des conditions d'environnements sonores bruyants.

Les agents du centre d'appel doivent avoir conscience de ce fait pour prodiguer des conseils pour les utilisateurs ayant des difficultés pour s'authentifier.

Où sont stockées les empreintes et quelle place occupent-elles ?

Les empreintes vocales sont simplement une représentation mathématique des caractéristiques uniques de la voix d'une personne. Elles sont stockées dans un emplacement dédié et sécurisé sur un serveur. Une empreinte vocale n'occupe que 20 ko d'espace disque. Une application avec 1 million d'utilisateurs ne nécessite donc pas plus de 20 Go de stockage.

Quelqu'un peut-il dérober mon empreinte vocale et obtenir l'autorisation d'accès au système en l'utilisant ?

Les empreintes vocales sont juste une matrice de chiffres représentant les caractéristiques physiques de la voix de la personne. La CNIL parle de gabarit. Ces informations ne peuvent pas être utilisées pour accéder directement à un système. Il n'est également pas possible de "reconstituer" un enregistrement factice avec votre voix sur la base de ces empreintes vocales.

Quelles sont les réglementations applicables à la biométrie vocale ?

En France, la CNIL indique que pour tous les types d'usage (contrôle d'accès Salariés/visiteurs, contrôle d'accès clients ou autres) le dispositif doit faire l'objet d'une demande d'autorisation.

La biométrie vocale est-elle conforme aux exigences de la CNIL concernant les dispositifs de contrôles biométriques ? Fiche pratique CNIL

Oui, cette technologie est connue de la CNIL au même titre que de nombreux autres dispositifs biométriques. La Cnil opère une distinction entre dispositifs qualifiés de "biométrie à trace" ou de "biométrie sans trace". La voix fait partie des dispositifs biométriques dits "intermédiaires", comme l'iris de l'œil et la forme du visage, c'est-à-dire qu'ils portent sur des caractéristiques qui ne peuvent pas ou très difficilement être reproduites à l'identique, contrairement aux empreintes digitales, que chacun laisse partout à son insu, et qui, avec des techniques modernes, peuvent être reproduites.

Nos solutions de biométrie vocales entrent également dans la catégorie des dispositifs à "Stockage sur un support non individuel".

En particulier, la plateforme de Nuance offre tous les outils d'administration en standard permettant simplement d'exercer les droits d'accès, de rectification, d'opposition et de révocation des empreintes.

Quel est aujourd'hui le sentiment des utilisateurs quant à l'utilisation de cette technologie de vérification d'identité ? Sont-ils prêts à donner et utiliser leur empreinte vocale à la place ou en plus de leur mot de passe ?

Nuance dispose déjà de plusieurs références client, où le nombre d'utilisateurs dépasse les 3 millions. Ces utilisateurs se sont enrôlés volontairement dans le système de biométrie vocal.

De plus, Nuance a récemment organisé une enquête Twitter pour déterminer le niveau de maturité du marché, et sonder des utilisateurs potentiels quant à leur opinion sur le sujet.

Cette étude montre à quel point il est devenu compliqué pour les utilisateurs de gérer leurs multiples identités numériques, et de se souvenir de tous les logins et mots de passe associés. (Les deux tiers détiennent plus de 10 mots de passes, et 19% plus de 30. L'étude montre également que, lorsqu'ils en ont la possibilité, les utilisateurs (80%) préfèrent réutiliser le plus souvent un même mot de passe pour différents services, introduisant un risque de sécurité important.

Parmi les résultats de cette étude, 77 % des sondés déclarent qu'ils préféreraient utiliser la biométrie vocale plutôt que d'autres méthodes si celle-ci offre une meilleure sécurité.

89 % de ceux-ci préféreraient utiliser leur voix (61%) ou la combinaison de la voix et d'un mot de passe (28%) plutôt qu'un mot de passe ou code PIN seul (11%).