Genesys Cloud Onboarding
ValidSoft Voice Verity

ValidSoft

30 Mooregate
London
EC2R 6JJ

United Kingdom

Tel.: +44 20 7164 6460

www.validsoft.com

vs-enquiries@validsoft.com

Confidentiality and Disclaimer

This document may contain references to information that has been obtained from sources believed to be reliable. ValidSoft does not guarantee the accuracy, completeness or adequacy of such information, and shall have no liability for errors, omissions or inadequacies. The recipient assumes sole responsibility for the interpretation and use of this material for its intended results. Predictions and forward-looking statements in this document reflect current expectations concerning future events and are subject to risks and uncertainties, many of which are beyond the control of ValidSoft. ValidSoft undertakes no obligations to update these statements as a result of new information. Opinions expressed in this document are subject to change without notice.

Introduction

This document is intended to help you understand how to use ValidSoft's Voice Verity solution within a CCaaS environment. It begins with a brief explanation of how synthetic speech detection functions and how conceptually it can be used, followed by a more comprehensive and practical description of how to incorporate it into your call flows.

Prerequisites

The ValidSoft solution requires setting up Roles, Queues, and Call Flows on Genesys Cloud CX before users can use it. This section is geared toward a Contact Center administrator since it involves common contact center terminology. If you think your Genesys Contact Center is set up correctly, please skip this section and resume to the "Onboarding process" section.

Note: This document assumes your organization has created Subscriptions, Divisions, and DID Numbers. If you need help with them, please consult your Genesys representative.

Roles and Permissions

Users on Genesys should have a specific role to be able to make and receive phone calls. This role is called "**PureCloud User**".

If your agents do not have this role, add the role via the following steps:

1) Click **Admin**.
2) Under **People and Permissions**, click **People**.
3) Find a user using the search or by paging through the list.
4) Click **More,** and from the menu that appears, click **Edit Person**. The view opens to the Roles tab.
5) Under View, select All.
6) In the Search box, begin typing the first few letters of the role you want to add and select it from the list.
7) In the Assigned column, click to enable the role for the user.
8) Repeat steps 6–8 to assign additional roles.
9) Click **Save**.

Phone Management

Agents on the Genesys PureCloud platform require a softphone to be able to receive/make phone calls. If your agent does not have a phone, you need to add a new phone for them via the following steps:

Configure a phone

1. Click **Admin**.

2. Under Telephony, click **Phone Management**.
3. Click the **Phones** tab.
4. Click **Add Phone**.

Phone tab

Configure settings on the **Phone** tab.

1. Type a name in the **Phone Name** box.
2. From the **Base Settings** list, select the base setting configuration you created.
3. From the **Site** list, select your site.
4. From the **Person** list, find the user you want to assign the phone to.
5. If you need to configure additional settings to your phone, follow the **Phone Configuration**.
6. Click **Save Phone**.

Queues

Queues are the “waiting lines” of interactions. Agents select the “On Queue” status to enter their predefined queues. Your Contact Center should have at least one queue for voice channels with agents in it to be able to receive interactions and use ValidSoft Voice Verity.

Create a new queue

If you don’t have a queue, create one by following these steps:

1. Click **Admin**.
2. Under **Contact Center**, click **Queues**. The Manage Queues page opens.
3. Click **Create Queue**.
4. In the **Name** box, type a name that is unique to the queue.
5. Click the **Division** list. To locate the required division, scroll through the list of available divisions and select the appropriate division from the list.
6. To copy the configuration and membership from an existing queue, under **Copy settings and members** from, search for and select an existing queue.
7. Click **Save**. The queue configuration opens to the **General** tab.
8. Click the **Voice** tab.
9. Toggle the switch to On for **Voice Transcription**.
10. Click **Save**.

Assign members to a queue

If you need to add yourself as an agent to your queue, follow the below steps to add members to your queue:

1. Click the **Members** tab.
2. To add a member to the queue, click **Add User**.
3. To search for users by a filter other than text, click the **Text** list and select from one of these filters:

- Division
- Group
- Language
- Location
- Reports To
- Skill

4. In the **Enter a value** box, begin typing the contact's name.
5. Select the appropriate match from the results.
6. To add the member or members to the queue, click **Add Selected**.
7. To remove a member from the queue, under **Action,** click **X** next to the member's name and then click **Confirm**.
8. Repeat steps 2–6 to add more members to the queue.

Architect

ValidSoft Voice Verity will require at least one Inbound Call Flow in **Architect** to use the **Queue** created in the above section and send the inbound call to the agents within the queue.

ValidSoft may provide you with a sample call flow that you can use. Follow the steps below to import this call flow into your **Architect** and configure it to complete the process. Alternatively, you can use your own flow.

Note: Importing a flow from an outside organization may require configuration updates, even for items such as data actions with similar names. Until you update any configuration issues, the Architect returns validation errors in the flow.

Importing a Flow

The **Import** command imports a flow configuration. Importing a flow will not merge with the existing configuration. When you import a flow configuration file into another flow, the **Architect** overwrites any existing configuration in the original flow upon import. To import a flow:

1. From the **Architect** home page, click or hover over the Flows menu and select the desired flow type.
2. Click the +Add button to create a new flow.
    - In the **Name** box, type a name that is unique to the flow
    - In the **Description** box, type a friendly description for the flow
    - Click the **Default Language** list. To locate the required language, scroll through the list of available languages and select the appropriate language from the list
    - Click the **Division** list. To locate the required division, scroll through the list of available divisions and select the appropriate division from the list
3. Click **Create Flow**

4. Click the arrow at the end of the **Save** menu and select **Import**. The Import a flow dialog box opens.
5. Download the configuration file onto your computer using the flow provided by ValidSoft.
6. Click the **Select a Configuration file to import** link and navigate to the .i3flow configuration file you want to import.
7. Click the **Import** button.
8. Save the flow.

Configuration of a Flow

To modify the queue to be used to route the call:

1. Click “Transfer to ACD” data action menu
2. Click the Queue selection menu and select the Queue you have created

Once you have completed the configuration, **Save** and **Publish** your Inbound Call Flow.

Using the Architect Flow

Add a call routing configuration entry to associate a telephone number or number with a call route. Callers will be routed to the appropriate call flow when they dial a number specified in the call routing configuration. This document assumes your organization has already provided Genesys with DIDs (call numbers).

Create call route

1. Click **Admin**.
2. Under **Routing**, click **Call Routing**.
3. Click **Add Call Route**.
4. In the **Name** field, type a unique call routing name. This name appears in your list of entries on the Call Routing page.
5. In the **Division** field, enter the division of the call route. This name appears in your list of entries on the Call Routing page.
6. To add phone numbers to associate with the configuration, do the following:
    a. From the **Inbound Numbers** drop-down list and select the required inbound numbers.
    b. Begin typing the telephone number string in numeric format only. When the number appears in the list, select the check box beside the number. Already assigned numbers are indicated as such, including the identification and type of assignee.
    c. Repeat step 6b to add more telephone numbers
7. Under the **What call flow should be used?** click the list and select the call flow you published in **Architect** from the list.
8. Click **Create**.

Onboarding process

The ValidSoft Voice Verity solution is available as a widget installed and accessed via Genesys AppFoundry. Enabling this for your environment is a prerequisite to any use of the solution. The process for onboarding is described in the following section.

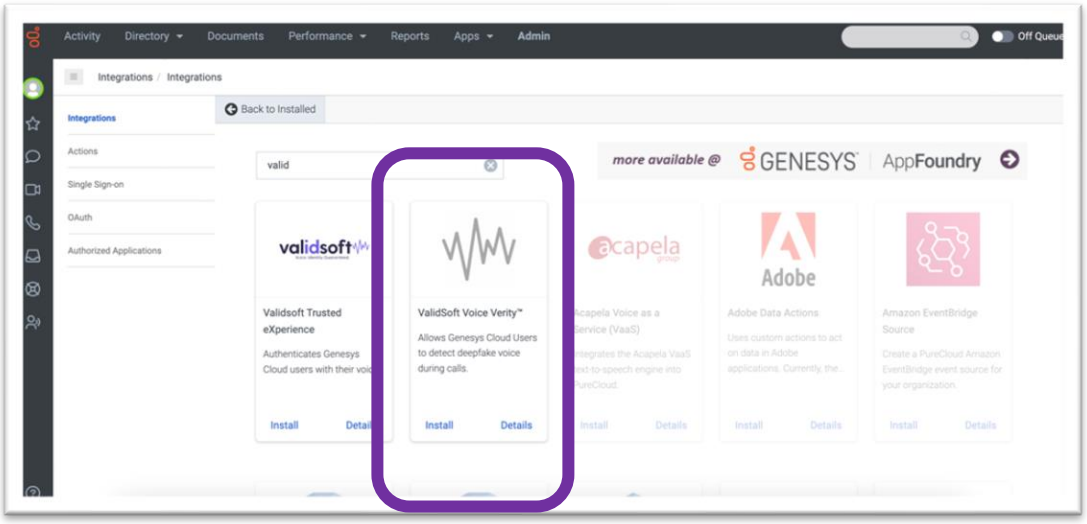
What is this?

ValidSoft Voice Verity is a deepfake detection solution that allows your organization to monitor incoming calls and notifies your agents if there is a threat of synthetically created audio being used by the calling party.

The deepfake detection will run in the background while the callers are speaking to agents and will notify the agent of the output of the deepfake audio detection via the widget offered by ValidSoft.

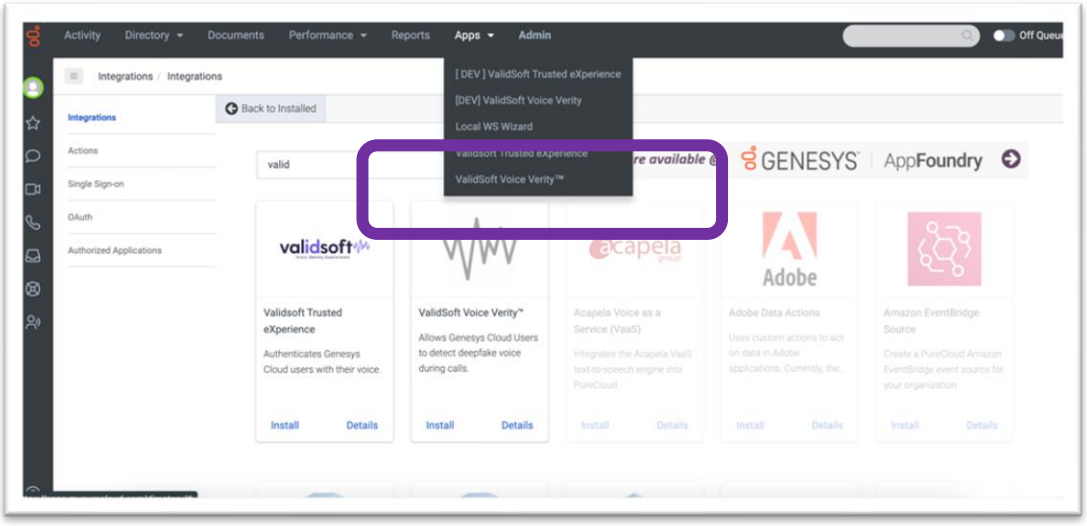
How to set up?

1. Click **Admin**
2. Under **Integrations**, click **Integrations**
3. Click the **+Integrations** button in the top right corner
4. Start typing a few letters of “ValidSoft Voice Verity,” and you will see the ValidSoft app in the search results:

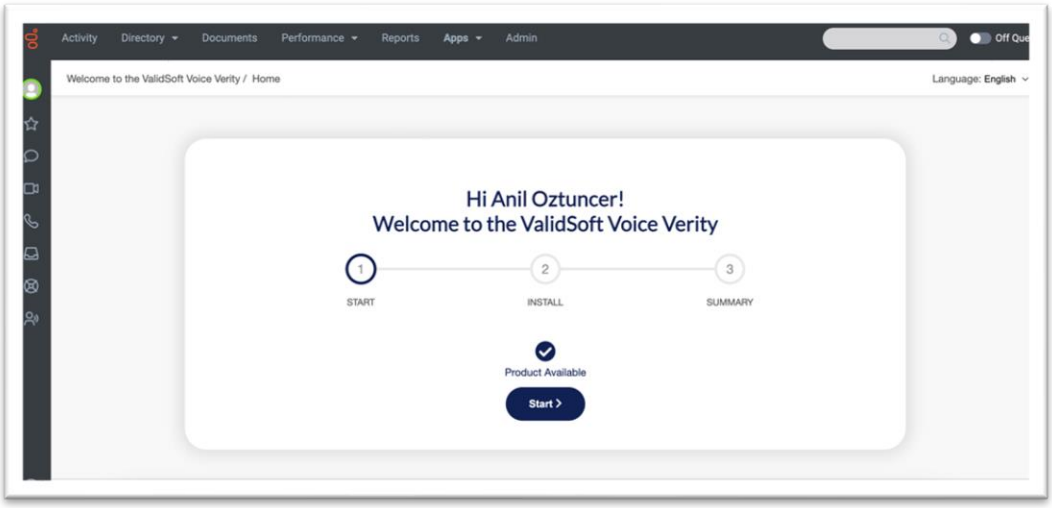


5. Click the **Install** button to install the app. Then, provide access to the requested permissions and agree to the Terms of Service. The ValidSoft Voice Verity view will open.
6. Under the **Details** tab, toggle the "Inactive" switch to activate the app.
7. Click Yes to the confirmation question.
8. Click **Save**.

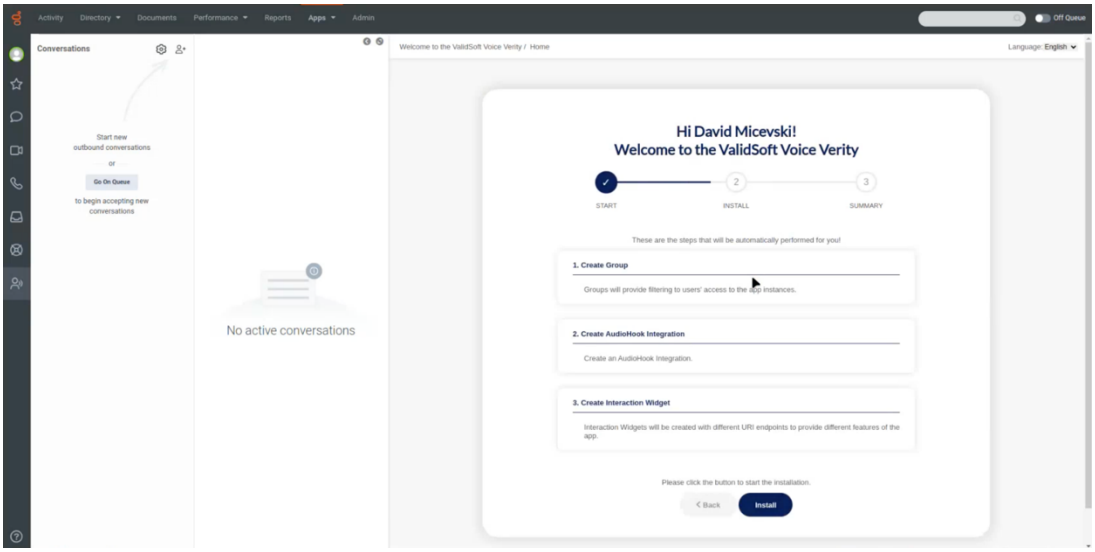
9. **Reload the page.** After the page is reloaded, the application will be available in the **Apps** menu:



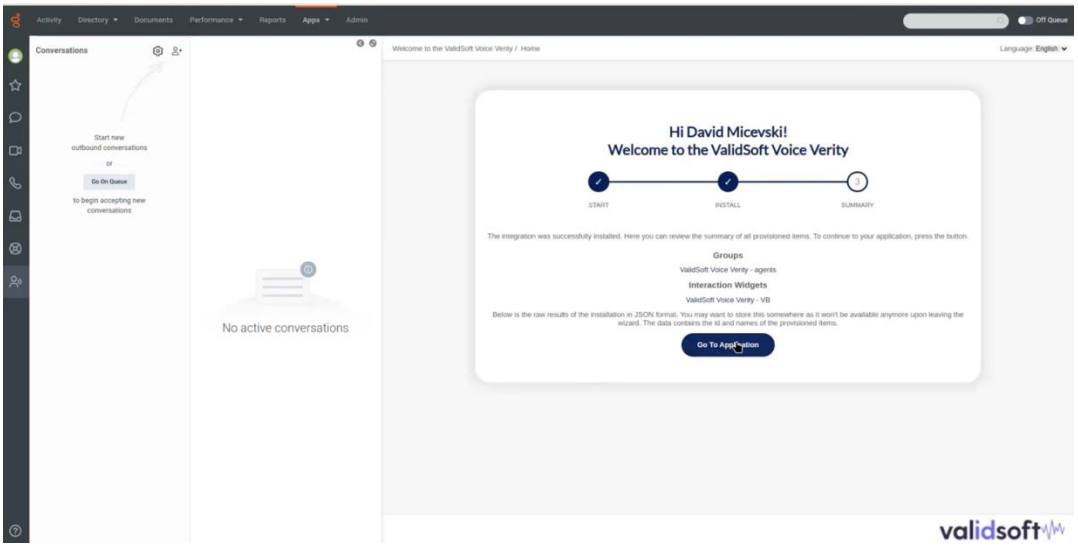
10. Click the **ValidSoft Voice Verity** application in the **Apps** menu. The installation wizard home page will display the following page:



11. Click **Start**.
12. The summary of the installation steps will be displayed:



13. Click **Install**.
14. The summary page will be displayed upon successful installation:



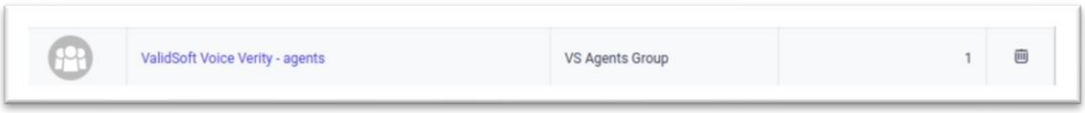
If you want to visit the configuration page for each individual component created by the install wizard, click the respective links under Groups, Interaction Widgets, and Audiohooks.

15. After the Installation Wizard finishes its installation successfully, the following components will be created:
    - **ValidSoft Voice Verity** Premium App. This is available in the Apps menu, which currently serves to display this documentation.
    - **ValidSoft Voice Verity** Interaction Widget. This visual component is available to an agent during the call and allows monitoring of deepfake detection capabilities.
    - **ValidSoft Voice Verity—Audiohook—Genesys AudioHook allows audio to be fed into the Voice Verity engine without** any configuration. More information on Audio Hooks can be found here.
    - **ValidSoft Voice Verity—agents Group. The Installation Wizard creates this group by default and assigns it** access to the interaction widget.

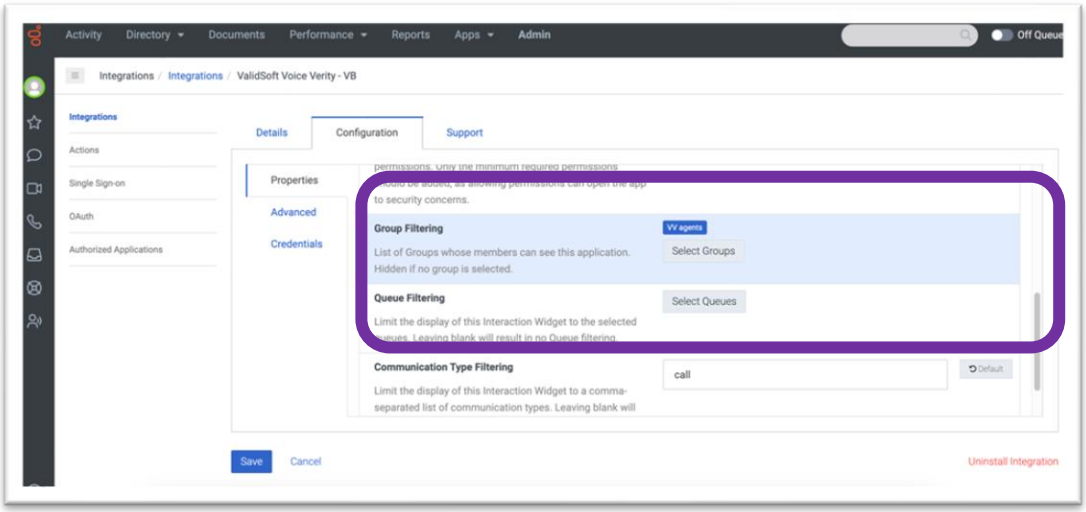
Configuration

There are two ways to enable agents from your organization to use ValidSoft Voice Verity:

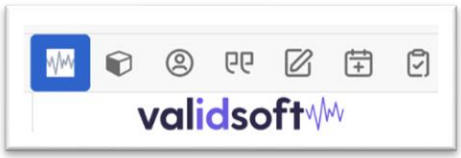
1. Assign eligible agents to a **ValidSoft Voice Verity - agents** group created by Installation Wizard:



2. Change group in Group Filtering settings at **ValidSoft Voice Verity - Widget** Interaction Widget -> Configuration -> Properties -> Group Filtering



All members of this group will have access to the **Trusted Voice Verity Widget** during calls:



Further configuration is not required.

Deepfake – Some Background

Deepfake audio (also referred to as Synthetic speech) encompasses sounds generated by software, either through Text-to-Speech (TTS, which produces an audio output sounding like a pre-trained speaker) or Voice Conversion (VC, in which existing audio is transformed or modified to mimic a particular speaker's voice).

Deepfake audio generation is typically done with Deep Neural Network (DNN) technology, a form of Artificial Intelligence (AI) that can be used to create audio recordings that sound like a specific person saying words or phrases they never actually said. While such systems have been available for decades, recent advances in generative AI make such technology much more accessible. Furthermore, they are able to create realistic, cloned speech with ever shorter audio from the targeted speaker required to train the underlying model. It has recently become an emerging method for perpetrating fraud and identity theft. Recently publicized examples include using deepfake technology, whereby security professionals "fooled" their own banks' voice biometrics solution or agent handling the calls.

Attack Vectors and Risks

With the availability of deepfake-generating software tools, the technology has been brought within reach of fraudsters and cybercriminals, and as always, they will devise novel ways of deploying it for monetary gain.

Voice deepfakes, in particular, have become a popular method for perpetrating fraud and identity theft. Recent publicity has focused on examples using deepfake technology from ElevenLabs, an example of a person fooling his own bank's voice biometric solution, and of the actress Emma Watson, who appears to be reading a copy of Mein Kampf. According to a Wall Street Journal article, the first published nefarious use of deepfake audio occurred in 2019.

According to the Wall Street Journal, the CEO of an unnamed UK firm received a phone call from supposedly the CEO of the firm's German parent company, requesting an urgent transfer of €220,000 ($243,000) to a Hungarian supplier. As the UK CEO recognized the German CEO's voice, including "the hint of a German accent and the same melody", he complied with the request. The money was subsequently moved from Hungary to a Mexican account and further disbursed. The details of the attack, but not the company, were shared with the WSJ by the company's insurer.

Apart from confirming that this form of deepfake attack is now in the wild and no longer just theoretical, it also confirms that the fraudsters are using synthetic "voice skin" technology, as distinct from creating static deepfake recordings. A synthetic voice skin allows the fraudster to speak in a conversational manner with the target, with the fraudster's voice being converted by the "skin" to sound like the impersonated voice. Whereas previously, voice skins have been of lower quality than static deepfake recordings, the technology used in this incident was clearly good enough to fool a CEO who would easily recognize his colleague's voice.

As with any fraud vector that is shown to work, it will only become more common and be used in ever more original ways, and in the case of deepfakes, the technology will evolve, allowing them to become even more realistic.

Unified Communications and Omni-Channel strategies mean organizations, including banks, will increasingly communicate with their customers using browser-based video/audio, for instance. This could be with a human agent, but in the future, also Artificial Intelligence (AI) based agents.

Imagine, therefore, a video/audio conversation between a high net-worth client and their private banker. If the client looks and sounds authentic and, of course, can provide the answers to any security questions (as they invariably would), why would the banker not acquiesce to any instructions the client gives?

How to Address the Problem

The rise in power and effectiveness of Deepfake audio has been directly driven by the ever-more rapid advances in AI tools, especially in DNN (Deep Neural Networks) technology. However, the contrary is also true; the ever-improving methods of identifying people from their voices (Voice Biometrics) have also hugely benefited from these same advances. We are a leading supplier of Voice Biometric solutions, and we have used these advances to improve the detection of real users. We have also used the same advances to detect the presence of deepfakes.

This Deepfake detection solution is integral to our solution, as are other anti-spoof techniques; all use DNN techniques to detect the presence of DNN-created audio, a real case of the same tools being used by both the hunters and the hunted.

As the threat from Deepfakes has become clearer and is starting to impact all walks of life, we have chosen to offer our Deepfake detection solution as a standalone solution, supporting both cloud and on-premise deployments and any channel that supports audio. More importantly, it is not a biometric solution, so it requires no enrollment, no consent, no PII implications, and no data holding.

Practical issues

Building or modifying the call flow

Voice Verity Deepfake detection can be applied as a standalone solution or with ValidSoft's Voice Biometrics solution. In both cases, the agent can be presented with a graphical and numerical result.

However, in our view, it is better if a synthetic audio caller can be detected before they arrive at the agent. In this way, the caller can immediately take the 'red route' towards trained operators, a fraud handling team or whatever your preferred method is for handling this type of fraudulent call. We suggest that the Deepfake detection component should be monitoring your callers from the moment that audio is available, especially if you use an IVR to extract user details such as account numbers. Deepfake detection can remain in the call flow to include the agent connect time so that a switch from real to fake audio can be detected, as well as highlighting any uncertain audio to the agent.

Because Deepfake detection is non-intrusive, any action that needs to be taken, e.g., detecting fake audio, can be initiated either automatically or by the agent, whichever combination is most appropriate for your use case.

Out of the box

Deepfake detection relies on monitoring the audio stream to detect artifacts and anomalies that are not normally present in real speech. The out-of-the-box solution is set up to detect under a generic set of conditions, which may be sufficient for your particular use case. However, it is important to consider sensitivity to the local acoustic conditions and the need for adaptation.

Tuning

Best practice recommends testing on audio as close as possible to the targeted use scenario (e.g. if targeting telephony on a platform with some given codec, use some real-world audio with the same conditions when testing) and analyzing results with the soft decision (the probability score given by the detector).

In some cases, the out-of-the-box solution benefits from some additional tuning to focus on more specific local conditions. In most cases, real and fake audio has a clear separation between the two streams. However, local acoustic conditions can move the null point between real and fake. This gives rise to the need to adjust the detection and alerting thresholds by tuning them using real-life data to be analyzed and for conclusions to be drawn. We recommend that it is not applied until several months of real volume activity have taken place. If you feel that your voice biometric solution would be improved by tuning, please request further information.

Results

The widget will present the agent with a 'Risk Bar', which is colored pale orange for a situation where there is a possibility that synthetic is present and red for the cases where there is no doubt that it is present. The actual score is embedded in the bar and can be exported elsewhere into your business flow.

Agent training

Agents will need training to understand the results that are returned by the Deepfake detection tool. In most cases, this will be a simple explanation of the two sections of the 'Risk Bar' and the actions that they should take when they occur.

Consent

Voice Biometrics relies on the inherently personal characteristics of the human voice, and therefore, the associated data and processes are often considered a special class of personal information. In many jurisdictions, additional regulations apply to their use and management.

The most common condition for the use of Voice Biometric data is user consent. This means that the user is aware and gives permission for you to carry out the processing in advance of any taking place.

However, because we neither analyse the audio content nor who the audio's owner is, there are no issues with privacy. In other words, we do not record, analyse, or retain any PII, and there are no requirements for obtaining consent when the Deepfake detection solution is used in a stand-alone mode. If it is used with our Voice Biometric solution, consent is required for that part alone.

Indirect (Genesys) Costs

ValidSoft's Deepfake detection tool relies on monitoring the audio stream, however it is important to note that this is a chargeable component by Genesys (aside from and in addition to ValidSoft's charges for use of the app. To limit this per call charge to a small percentage of the overall charge made by Genesys for audio, ValidSoft has set several default values for the duration that the tool operates. This is set at 60 seconds of audio (whether speech is present or not) after which the tool will cease consuming the audio stream and hence stop any Genesys based charges. The agent can restart this process for an additional 60 seconds using the button on the widget if required. Additionally, if your organisation prefers to have a different analysis duration, ValidSoft will be pleased to adjust this value for you.

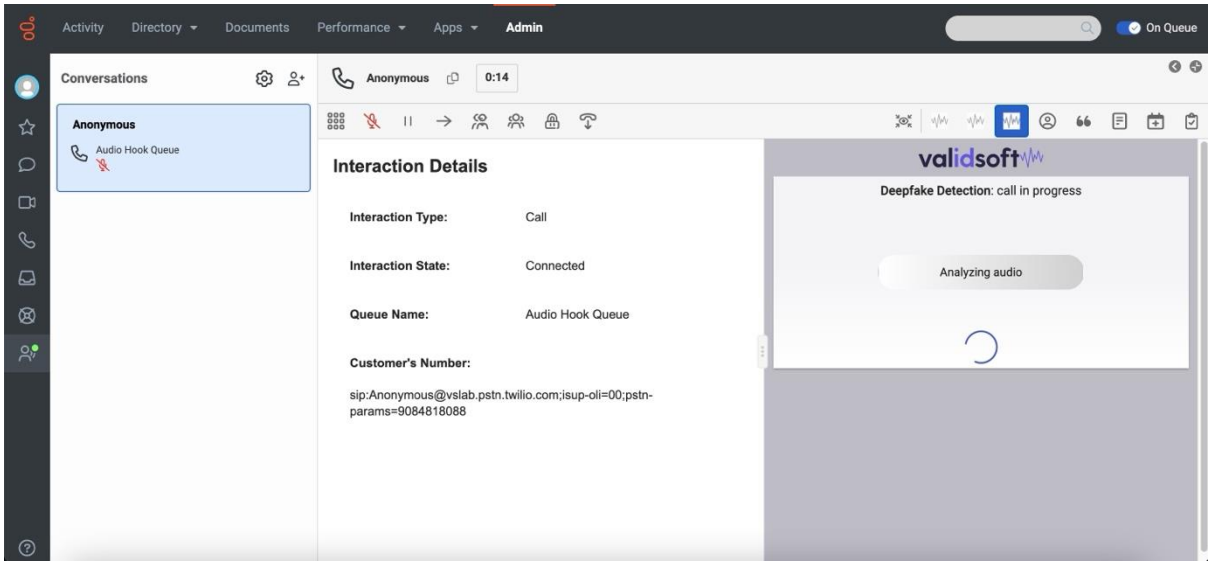
Use of the Widget on Agent Screens

The following screenshots show how the widget will appear to agents and the flow sequence that they will need to follow.

Inbound Call Presentation – Initial Launch

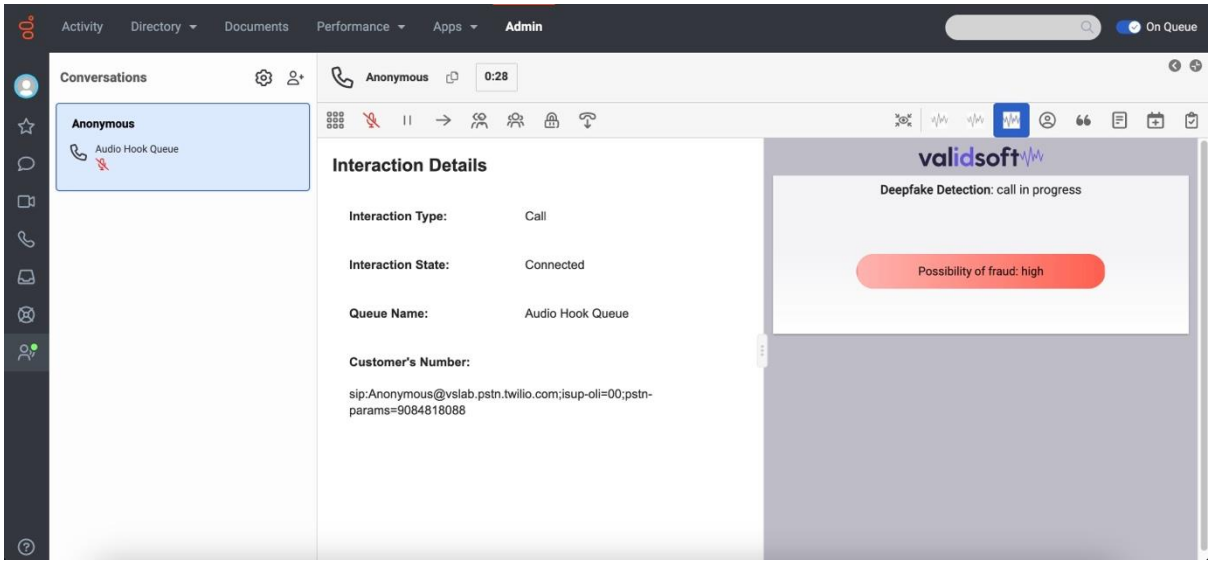
When the agents first launch the widget, they will be greeted with the below screen. This means the Voice Verity service is analyzing the customer audio to determine whether it's a genuine person speaking or a deepfake/synthetically created audio.

Depending on the caller's speech pattern, agents usually wait 4-5 seconds for the service to determine the source of the audio.



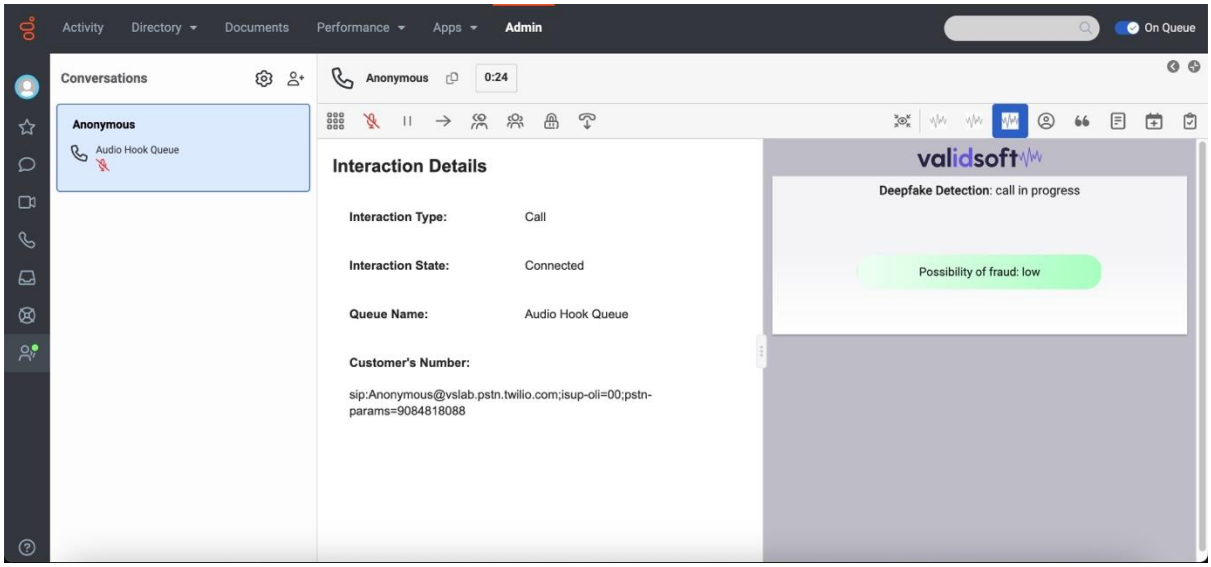
Deepfake Detection

When the tool detects that the caller might be using synthetically created audio, agents see the screen below. This shows them that the probability of fraud is high, and they need to proceed with caution.



Genuine Audio Detection

When the caller does not use synthetically created audio and speaks typically, agents will see the screen below. The green flag means the probability of the caller using a deepfake attack is low, and agents can usually resume their operation without heightened security.



Testing

Methodology

There are two parts to the testing of any Voice Biometrics solution that need to be considered, even when simply searching for Deepfakes.

- Functional Testing
  This addresses the basic processes for detecting synthetic audio presented by the Deepfake detection tool.
- End-to-End and Performance Testing
  This tests the system's mechanisms within the larger ecosphere and ensures that the Deepfake detection solution's performance meets the expectations for detecting real and fake audio in a production environment.

Functional Testing

The following considerations should be applied to support functional testing of the Voice Biometrics capabilities.

The amount of original speaker audio used to train the underlying synthetic generation engine can have an impact on the detection rate. As the amount of original speaker audio used increases, the deepfake detection accuracy tends to be lower since the synthetic generation engine has more data and produces a better synthetic result. Nonetheless, certain traits and easily recognizable artifacts (such as a slightly robotic sound) will be present, whatever the amount of audio used in training.

When assessing the detection of Deepfake technology, you must take certain considerations into account as these not only alter ValidSoft's ability to detect acoustic anomalies, but it also impacts the ability to be orchestrated in "the real world".

The test corpus should be sufficient size to ensure that the population is able to give statistically meaningful result to the tests and any subsequent analysis of the results. If this isn't done, over generalization or "skewing" the results is possible due to the sample size either being too small or not reflective of the wider demographic for deployment.

To obtain a fair statistical representation of the performance of a given service it is necessary to:

(a) have a statistically relevant sample set of speakers and;

(b) ensure that the capture of True and False scores or sessions is segregated and documented to ensure outliers are not caused by unexpected usages of the services during the capture process.

Test actions:

- Present original audio and synthetically generated audio to the Voice Verity solution
- Observe the results taking into account the caveats above, especially those related to a sufficiently meaningful volume of test cases. ValidSoft are happy to advise on this point but you should expect to be at least 100s of different test audios

What to do when things don't go as you expect

Failure to Install or no Widget Visible

We sometimes see clients who have downloaded the app from AppFoundry and who do not see the widget on the agent desktop. Sometimes the entire onboarding process hangs and will not complete.

This is almost invariably caused by missing a setup or configuration component or carrying out one of the steps in a different order to that described in Onboarding Process at the start of this document.

Solution

1. De-install the app by *selecting and then choosing delete from the options.*
2. Ensure that roles and permissions are correctly set as described under the Prerequisites section at the start of this document
3. Ensure that Queues are setup and in place and that at least one member (agent) is assigned to a queue.
4. For test purposes, we recommend that you install the sample call flow described in the Architect section and configure it following the guidance in that section.
5. Follow the instructions in the Onboarding Process section in the order in which they are given.

If you see a pattern of failures from users, please contact us (cs@validsoft.com) and we will look more deeply into the potential causes of the failures.

About ValidSoft

ValidSoft offers the world's fastest, most accurate and precise voice biometrics authentication and identification platform: ValidSoft VIP®. It works in the Cloud, Private Cloud, SaaS, On-premise, and On-device and is a true Omni-Channel with no loss of performance. ValidSoft's VIP technology differentiates itself on three strategic pillars: Precision & Accuracy (Security), Data Privacy & Protection (Integrity) and Omni-Channel Excellence (Consistent Experience & Future Proofing).

ValidSoft saves businesses money, stops fraud, and eliminates consumer frustration by providing a seamless, omni-channel solution that enables secure, fast, friction-free customer engagement.

Our website can be found at **www.validsoft.com**.