LumenVox®
Speech Understood

PCI / HIPAA BEST PRACTICE RECOMMENDATIONS

Executive Summary:
Data security is at the forefront of every datacenter manager's mind. This paper describes best practices for securing sensitive data when working with LumenVox products, and outlines the extensive functionality that has been made available for protecting this data.

Audience:
This white paper is a technical document that is written for developers. It assumes some experience with speech technologies (including familiarity with basic speech recognition grammars and the Speech Semantic Markup Language for TTS).

INTRODUCTION

Whenever developing any application, not just applications using LumenVox technologies, it is important to be aware of potentially sensitive data that may be associated with the application, such as credit card information. This is of particular importance when working with PCI or HIPPA compliance.

By taking these things into consideration when developing applications at the core level, it is possible to identify and describe whether and where any sensitive data is.

LumenVox products are used in a large number of diverse applications, from simple demonstrations to integrated banking applications, and as such, there is the possibility that sensitive data may pass through LumenVox software. It is essentially impossible for the software to know whether any data should be considered sensitive or not, so the application developer has a responsibility to understand what the risks are and determine whether data being processed could be considered sensitive, and more importantly, take measures to limit exposure of that data.

Developers can utilize some of the built in features of LumenVox software to help control this exposure when needed.

Compliance Requirements

Customers often ask whether LumenVox software needs to be certified as PCI or HIPAA compliant. The current LumenVox ASR/TTS/CPA software suite is considered a back office system in the eyes of both PCI and HIPAA, so there really is no need or option for any compliance certification. These specifications place the burden of securing data on the integrators and users of datacenters utilizing this software. As such, LumenVox have implemented several security enhancements to the software, giving developers and users features that allow them to protect sensitive data in a number of ways. This document is designed to assist such developers and users in this task by clearly identifying areas of the software that may expose sensitive data and also some ways in which that data can be protected.

HOW CAN LUMENVOX HELP

This document was formulated in order to provide our users that are concerned with PCI and/or HIPAA compliance with information that may be useful when performing a security evaluation relating to the use of LumenVox software and to assist in protecting potentially sensitive data. This document achieves this in the following ways:

1. **Identify areas of potential risk**
   Throughout the remainder of this document, various areas of potential risk are identified, allowing users, application designers and IT managers to better understand where potentially sensitive data may be exposed when using LumenVox software.

2. **Provide details of any risk mitigation measures that can be taken**
   Along with each potential risk that is identified, measures that can be taken to limit or prevent exposure of potentially sensitive data are also described.

Our hope is that this document assists users, developers and IT managers understand and best manage the risks associated with potentially sensitive data relating to PCI and/or HIPAA compliance. These same issues may also be of interest to system administrators reviewing data security in general.

You should check to see if a newer version of this document is available before acting upon the information it contains. In addition, it is the responsibility of the users, application developers and IT managers deploying and using LumenVox software that need to perform their own due diligence and use the information contained in this document as best practice guidance only.

LumenVox, working in conjunction with some of our major clients and partners have developed a number of features that assist in securing or suppressing sensitive data.

In general, if the application developer knows that data could be potentially sensitive for an ASR or TTS interaction, the LumenVox software can be configured to enable various security features to protect that sensitive data as needed. This is described in more specific detail for each of the areas of potential risk that are described later in this document.

Configuration settings can be specified at startup that enforce sensitive data protection functionality by default, and also disable the logging or recording of events as well as suppressing potentially sensitive data. This level of protection may be too limiting in certain situations, since it may prevent diagnostic reporting, but being aware of this option is important if your primary goal is data security.

DATA DISCOVERY

Determining where potentially sensitive data may be located is crucial when analyzing security risks associated with any software. LumenVox are continuously working internally and with partners to improve security and identify areas of concern. Specifically, a number of areas have been identified where potentially sensitive data may be found, which this guide will attempt to describe in as clear a manner as possible in the following sections.

Note that in most respects, the CPA functionality may be considered an ASR resource for the purposes of this document, so references to ASR resources also covers the CPA resource type as well.

RTP Stream Data (ASR Resources)

As with any other network communication traffic, when receiving DTMF data over an RTP stream, the RTP-Events themselves may contain the numbers of a credit card or other sensitive information when viewed or recorded with a network protocol analyzer utility.

In addition to DTMF events, voice audio is generally used with ASR resources which may be contained within a similar or the same RTP stream and may similarly contain the numbers of a credit cards or other sensitive information when accessed with a network protocol analyzer utility.

> **Risk Mitigation:**
> Because data is being sent in an unencrypted manner across the network, LumenVox software cannot prevent such exposure. In future versions of the software, there will be support for encrypted network traffic, which will include audio/DTMF (RTP) streams as well as session control streams (RTSP/SIP/MRCP).

RTP Stream Data (TTS Resources)

TTS audio being sent to an RTP stream from LumenVox software may contain the numbers of a credit card or other sensitive information when accessed with a network protocol analyzer utility.

There are no DTMF-Events generated by the LumenVox software that are sent to an RTP Stream.

> **Risk Mitigation:**
> Because data is being sent in an unencrypted manner across the network, LumenVox software cannot prevent such exposure. In future versions of the software, there will be support for encrypted network traffic, which will include audio/DTMF (RTP) streams as well as session control streams (RTSP/SIP/MRCP).

Grammar Data

Grammar files are often being fetched across network resources, which may also contain potentially sensitive information.

> **Risk Mitigation:**
> As LumenVox API and MRCP functions are extremely flexible in the URI references that may be used to specify a grammar file location, this may be a file system URI, a relative path (from the client), or a remote HTTP/HTTPS URI. Assigning special folder permissions to these locations may prevent unwanted access; however this does not protect the data as it is being fetched by the application. Storing sensitive grammars on a HTTPS server will offer some protection as it is traversing the network. In addition, LumenVox grammars can be compiled into a proprietary binary format, which can also be accessed locally or remotely across a network in the same way as uncompiled/raw grammars. This type of pre-compiled grammar is not human-readable since it is in a binary format, and offers additional protection for both security reasons as well as to protect important IP built into some grammars. See the knowledgebase article on using precompiled grammars for more details.

SSML Data

Text to be synthesized for a TTS resource may be either plain text or a markup language (SSML) used to express more clearly the way in which things should be pronounced.

Many of the LumenVox API and MRCP functions allow users to specify URIs to use when fetching SSML data being used for speech synthesis using a TTS resource. This SSML data may contain potentially sensitive data, and since this is outside of the LumenVox code, it may be viewed using network analysis tools.

> **Risk Mitigation:**
> Storing this SSML data (files containing the desired SSML content) on web servers utilizing a secure protocol such as HTTPS may be used to protect the data as it traverses the network.

Log Data (API Users)

Logging information is recorded to a number of file-based log files that are stored in plain-text format. This logging is similar for all LumenVox APIs, so that users of our C and C++ interfaces to ASR, TTS and CPA resources can be described in a general sense from a security perspective. The LumenVox knowledgebase contains a Logs Overview article describing which log files are used and where to locate those log files.

Logging within LumenVox software can be configured for various levels of verbosity. The default verbosity level is 1, which records minimal information relating to any errors or warnings encountered. Setting higher levels of verbosity will log more information, including diagnostic information at higher verbosity settings.

Since these logs are recording activity within various parts of LumenVox software, it is important to understand that sensitive information may be contained within these logs, such as the result of ASR recognition, or the raw text to be synthesized by a TTS resource.

Risk Mitigation:
There are a number of things that can be done to prevent potentially sensitive data to be exposed to the log files. One option may be to assign special permissions to the logs folder, preventing unauthorized access to these files. Another option may be to limit the verbosity setting to a minimal setting (1), which will only record errors and warnings, which do not contain any sensitive data.

In addition to these simple options, there are specific API functions that are designed to force a specific ASR or TTS resource into a "secure_context" mode, where all potentially sensitive information will be suppressed from the logs, replacing potentially sensitive data simply with the word "_SUPPRESSED".

When dealing with the ASR resource, it can be tasked to process both speech and DTMF requests, and when operating in secure_context mode, these DTMF events will be suppressed in the same fashion. Also note that each ASR and TTS resource (port) each have their own separate and independent settings for this secure_conext mode.

There are also configuration settings available allowing this secure mode to be on by default at startup if that is desired. More details on controlling and using this mechanism is described in these knowledgebase articles for ASR and TTS resources respectively and the configuration settings are described in the client_property.conf article.

Log Data (Media Server Users)

The LumenVox Media Server can be considered a wrapper around the C/C++ APIs for both ASR and TTS resources. In addition to the API functionality, the Media Server provides additional connectivity to support MRCPv1 and MRCPv2 clients and their respective media streams and control sessions. This service therefore uses many of the same API logs described above, and uses the same secure_context settings in addition to some others specific to MRCP.

The Media Server has its own set of additional log files that are created when operating. Verbosity of these logs can be controlled by the same configuration settings as the API.

In addition to API logging information, the Media Server can log inbound and outbound SIP/RTSP/MRCP/RTP-Event traffic, depending on its logging verbosity setting. If configured for minimal verbosity, only errors and warnings will be recorded, however with more verbose settings, the information passed over the network within the SIP/RTPS/MRCP/RTP-Event traffic may also be logged, which may contain potentially sensitive data.

Some specific areas where sensitive data may be logged include the logging of RTP-Events corresponding to DTMF keys within media_server_app.txt, which can be used to expose potentially sensitive information. Similarly, the logging of a recognition results from either speech or DTMF activity in media_server_app.txt may also expose potentially sensitive data.

Risk Mitigation:
To avoid logging sensitive data, select a minimal verbosity setting to ensure that only errors and warning messages are logged, which should not pose a security risk.

At higher verbosity settings, the secure_context configuration settings may be set to suppress logging of potentially sensitive data, in the same way as is described above for API (non-Media Server) users. Since users do not have direct access to the API functions allowing the modification of secure_context settings on a per-interaction basis, when running the Media Server, there is additional functionality provided to allow these settings to be controlled using Vendor-Specific-Parameters. See the Vendor Specific Parameters knowledgebase article for more details. Note that once again, there are separate and independent secure_context controls for both ASR and TTS resources.

MRCP Specific Audio Recording

MRCP save-waveform functionality maybe used to store (RTP) audio sent into the ASR resource. The Media Server can be configured to store these waveforms to files on disk, potentially exposing sensitive information via the files on disk.

Risk Mitigation:
The save-waveform functionality in MRCP needs to be explicitly configured and also activated in order to be used. This means that for potentially sensitive data to be exposed, the Media Server configuration needs to be set up to specify a target location for these waveforms. By default this is not set, which effectively disables the functionality. In addition, the MRCP application developer needs to activate the save-waveform session parameter for these files to be generated.

Also, once these files are generated, the RECOGNITION-COMPLETE message may report the Waveform-URI that can be used to access these files, which poses an additional potential data risk. In such cases, implementing access restrictions to these exposed URI's may also be desirable.

Disabling the save-waveform functionality for potentially sensitive interactions removes this risk, since these files would not be created.

The knowledgebase article on Media Server configuration settings can be used as a reference to disable or configure the waveform saving functionality; specifically the save_waveform and waveform_url_location settings. Basically, if the waveform_url_location setting in media_server.conf is set to a blank string, this will disable any possible save-waveform functionality within the MRCP interfaces in a way that ASR audio will not be recorded using that feature.

MRCP Specific Inline Grammars

Inline grammars that are sent to the recognizer resource may include sensitive data, which can appear in log files as well as in .callsre files.

Risk Mitigation:
Customers wishing to prevent this exposure can use non-inline LumenVox pre-compiled grammars, which are represented in binary form and cannot be easily read. If inline grammars are specified, and secure_context is enabled for the resource, the grammar will be suppressed from the logging files and also .callsre files. Another option may be to avoid using inline grammars in favor of grammar files stored in a secure (HTTPS) location, which may also be precompiled to further enhance security.

MRCP Specific Inline TTS

Inline SSML or plain text may be specified in certain MRCP requests, such as the SPEAK request. These contain potentially sensitive information, which can appear in log files as well as in .callsre files.

Risk Mitigation:
To prevent this exposure, users are encouraged to specify remote (HTTPS) locations to store such sensitive SSML files in favor of inline text to offer protection as they traverse the network. If inline SSMLs are specified and secure_context for the resource is enabled, the SSML will be suppressed from the logging files and also .callsre files.

MRCP Specific Inline Interpret-Text

The MRCP INTERPRE-TEXT request is used to request that the recognizer resource perform a text parse of the specified text. This text may contain potentially sensitive data, especially as it traverses the network.

Risk Mitigation:
Enabling the secure_context for the resource will prevent this text from being recorded in the log files and also .callsre files.

Tuning Data Diagnostic Logging

LumenVox software has the ability to generate diagnostic logging files, often called ".callsre" files because of their file suffix. These files can be used with the LumenVox Speech Tuner to analyze the performance of and tune speech applications.

These files can contain detailed information relating to ASR and TTS interactions, which may also contain potentially sensitive data, which may be exposed using the Speech Tuner.

Risk Mitigation:
Using the "SAVE_SOUND_FILES" configuration option, users can enable or disable this diagnostic logging feature, or control it using the PROP_EX_SAVE_SOUND_FILES option of the SetPropertyEx API functions. If this value is set to a non-zero setting, various aspects of ASR and TTS interactions are stored in these diagnostic logging files for possible later analysis. A value of 0 for these settings will disable the generation of these diagnostic logging files.

ASR interaction audio and recognition results will be suppressed from these diagnostic logging files if the secure_context option for the ASR resource is enabled.

TTS interaction audio and synthesized text will be suppressed from these diagnostic files if the secure_context option for the TTS resource is enabled.

In addition to these suppression options, recent versions of LumenVox software include an option to securely encrypt these diagnostic response (.callsre) files. When this option is used, any data stored in response files is encrypted using a public certificate/private key file combination specified by the user. An optional LumenVox public certificate may also be used if the data is intended to be shared with LumenVox.

Once this ASR/TTS interaction data is encrypted in such a way, the user must have access to the appropriate secure private key in order to be able to use the Speech Tuner to view that data.

An optional user-designated password can also be assigned to the private key when the certificate/key pair is created, to further enhance security, so that the user would need access to the hidden password/passphrase in addition to the appropriate secure private key file in order to access the data within the Speech Tuner.

SPEECH TUNER

The LumenVox Speech Tuner application by its nature is designed to analyze the raw data from ASR and TTS interactions in order to diagnose any performance issues as well as offer tuning options. This means that to perform this function, it may need access to potentially sensitive data, since those interactions also need to be tuned and diagnosed at times.

Data contained within the diagnostic logging (.callsre) files may contain potentially sensitive data, so anyone using or viewing the Speech Tuner may intentionally or unintentionally gain access to this potentially sensitive data.

Risk Mitigation:
If secure_context is enabled for ASR and TTS resources, the potentially sensitive information is suppressed from log files as mentioned earlier, but also from these diagnostic (.callsre) files. If the Speech Tuner is later used to view these suppressed interactions, the user will be able to see only basic information about the interaction, and not gain access to the potentially sensitive data.

Also, if encryption is enabled for interactions, any potentially sensitive data may be accessed in a secure way later if needed, as may be required for diagnostic or tuning purposes.

We recommend the use of encrypted diagnostic files, with optional password/passphrase protection to provide the maximum protection for sensitive data when needed. Using this method, whoever generates the secure certificate/private key pairs can also control access to those private keys, and therefore the data they protect.

DASHBOARD UTILITY

The Dashboard utility is provided with LumenVox product to allow browser-based control over settings and also provide access to logging and status information. Clearly this poses a security risk from an administrative perspective, and also a data security perspective.

Risk Mitigation:
There is an optional username/password combination that can be configured to restrict access to the Dashboard application. In addition to this, the application can be configured to communicate using the HTTPS protocol to protect traffic to and from the browser as it traverses the network.

If necessary, the HTTP/HTTPS functionality can be disabled entirely through configuration which prevents any remote connectivity to the Dashboard application, essentially disabling it.

ADDITIONAL ISSUES

In addition to the above areas that have been identified, there are a number of system specific concerns that involve physical computer security, which should also be controlled appropriately when dealing with secure or sensitive data. These topics are beyond the scope of this document.

We recommend consulting an experienced network systems administrator when dealing with any security issues.

LumenVox is a speech automation solutions company providing technology design, development, deployment, tuning and transcription services including the LumenVox Speech Recognizer, Text-to-Speech Engine, Call Progress Analysis, Speech Tuning Services, and SLM solutions. Based on industry standards, LumenVox's core Speech Software is certified as one of the most accurate, natural sounding, and reliable solutions in the industry.

For more information, visit www.lumenvox.com