



Hewlett Packard
Enterprise

IDOL Speech Server

Software Version: 11.5.0

Release Notes

Document Release Date: October 2017

Software Release Date: October 2017

Legal notices

Warranty

The only warranties for Hewlett Packard Enterprise Development LP products and services are set forth in the express warranty statements accompanying such products and services. Nothing herein should be construed as constituting an additional warranty. HPE shall not be liable for technical or editorial errors or omissions contained herein.

The information contained herein is subject to change without notice.

Restricted rights legend

Confidential computer software. Valid license from HPE required for possession, use or copying. Consistent with FAR 12.211 and 12.212, Commercial Computer Software, Computer Software Documentation, and Technical Data for Commercial Items are licensed to the U.S. Government under vendor's standard commercial license.

Copyright notice

© Copyright 2017 Hewlett Packard Enterprise Development LP

Trademark notices

Adobe™ is a trademark of Adobe Systems Incorporated.

Microsoft® and Windows® are U.S. registered trademarks of Microsoft Corporation.

UNIX® is a registered trademark of The Open Group.

This product includes an interface of the 'zlib' general purpose compression library, which is Copyright © 1995-2002 Jean-loup Gailly and Mark Adler.

Documentation updates

The title page of this document contains the following identifying information:

- Software Version number, which indicates the software version.
- Document Release Date, which changes each time the document is updated.
- Software Release Date, which indicates the release date of this version of the software.

To check for recent software updates, go to <https://downloads.autonomy.com/productDownloads.jsp>.

To verify that you are using the most recent edition of a document, go to <https://softwaresupport.hpe.com/group/softwaresupport/search-result?doctype=online help>.

This site requires that you register for an HPE Passport and sign in. To register for an HPE Passport ID, go to <https://hpp12.passport.hpe.com/hppcf/login.do>.

You will also receive updated or new editions if you subscribe to the appropriate product support service. Contact your HPE sales representative for details.

Support

Visit the HPE Software Support Online web site at <https://softwaresupport.hpe.com>.

This web site provides contact information and details about the products, services, and support that HPE Software offers.

HPE Software online support provides customer self-solve capabilities. It provides a fast and efficient way to access interactive technical support tools needed to manage your business. As a valued support customer, you can benefit by using the support web site to:

- Search for knowledge documents of interest
- Submit and track support cases and enhancement requests
- Access product documentation
- Manage support contracts
- Look up HPE support contacts
- Review information about available services
- Enter into discussions with other software customers
- Research and register for software training

Most of the support areas require that you register as an HPE Passport user and sign in. Many also require a support contract.

To register for an HPE Passport ID, go to <https://hpp12.passport.hpe.com/hppcf/login.do>.

To find more information about access levels, go to <https://softwaresupport.hpe.com/web/softwaresupport/access-levels>.

To check for recent software updates, go to <https://downloads.autonomy.com/productDownloads.jsp>.

Contents

New in this Release 5

Resolved Issues 9

Notes 10

Documentation 13

New in this Release

This section lists the enhancements to IDOL Speech Server version 11.5.0.

- New improved Neural Network technology is now used for Speech-to-text. You can use Speech Server version 11.5.0 with the new 9.0+ Language Packs to improve results accuracy. In addition, the new technology is typically faster than older server and language pack combinations.
- Speech Server speech to text tasks now generate word confidence values by default. The `stt` module `EnableConfidence` configuration parameter now has a default value of **True**.
- Language identification and Speaker identification tasks now normalize confidence scores to a value between 0.0 and 1.0 by default. You can modify this behavior by changing the `ScoreMode` parameter in the `ivscore` or `langid` module configuration to display the raw scores instead. This parameter is also available as a parameter for the `LangId` and `IvSpkId` standard tasks.
- The Speech Server standard tasks have been simplified so that you can use the same task for audio file and stream processing, with an `InputType` parameter to specify the correct input type (**file**, **data**, or **stream**). In addition, other task sets have been simplified. For all language identification tasks, you can now use the `LangId` task, with the `LidMode` parameter set to provide segmented, cumulative, or boundary mode identification.

As part of this change, existing tasks have been deprecated and replaced with the new equivalents.

The following table lists the deprecated tasks, and the new task to use instead.

Old Task (Deprecated)	New Task
WavToText	SpeechToText (with <code>InputType=File</code>)
StreamToText	SpeechToText (with <code>InputType=Stream</code>)
StreamToTextMusicFilter	SpeechToTextFilter
TelWavToText	SpeechToTextTelephony
ivSpkIdTrainWav	ivSpkIdTrainAudio (with <code>InputType=File</code>)
ivSpkIdTrainStream	ivSpkIdTrainAudio (with <code>InputType=Stream</code>)
ivSpkIdDevelWav	ivSpkIdDevelAudio (with <code>InputType=File</code>)
ivSpkIdDevelStream	ivSpkIdDevelAudio (with <code>InputType=Stream</code>)
ivSpkIdEvalWav	ivSpkId (with <code>InputType=File</code>)
ivSpkIdEvalStream	ivSpkId (with <code>InputType=Stream</code>)
ivSpkIdSetInfo	ivSpkIdInfo (with <code>TemplateSet</code>)
ivSpkIdTmpInfo	ivSpkIdInfo (with <code>TemplateFile</code>)
ivSpkIdSetEditThresh	ivSpkIdEditThresh (with <code>TemplateSet</code> and <code>TemplateName</code>)

ivSpkIdTmpEditThresh	ivSpkIdEditThresh (with TemplateFile)
LangIdSegLif	Language identification on feature files is deprecated. Use LangId.
LangIdCumLif	Language identification on feature files is deprecated. Use LangId.
LangIdBndLif	Language identification on feature files is deprecated. Use LangId.
LangIdSegWav	LangId (with InputType=File and LidMode=Segmented)
LangIdCumWav	LangId (with InputType=File and LidMode=Cumulative)
LangIdBndWav	LangId (with InputType=File and LidMode=Boundary)
LangIdSegStream	LangId (with InputType=Stream and LidMode=Segmented)
LangIdCumStream	LangId (with InputType=Stream and LidMode=Cumulative)
LangIdBndStream	LangId (with InputType=Stream and LidMode=Boundary)
afpAddTrackWav	afpAddTrack (with InputType=File and afpMode=standard)
afpAddTrackStream	afpAddTrack (with InputType=Stream and afpMode=standard)
afpMatchWav	afpMatch (with InputType=File and afpMode=standard)
afpMatchStream	afpMatch (with InputType=Stream and afpMode=standard)
afptAddTrackWav	afpAddTrack (with InputType=File and afpMode=robust)
afptAddTrackStream	afpAddTrack (with InputType=File and afpMode=robust)
afptMatchWav	afpMatch (with InputType=File and afpMode=robust)
afptMatchStream	afpMatch (with InputType=Stream and afpMode=robust)
afptRemoveTrack	afpRemoveTrack (with afpMode=robust)
afptDatabaseInfo	afpDatabaseInfo (with afpMode=robust)
wavPhraseSearch	PhraseSearch
wavToFmd	createFmd

- The new audio module has been added, which allows you to process audio files, binary data, or streams in the same module. This module replaces the `wav` and `stream` modules, which are now deprecated.

This module accepts the same parameters as the old `wav` module, and also a new `InputType` parameter, which defines what type of input to process.

The following tasks use the new module. As such, you can use these tasks for audio files or streams.

afpMatch	createFmd	phraseSearch
----------	-----------	--------------

afpAddTrack	dialToneIdentification	SNRCalculation
audioAnalysis	ivSpkId	speechSilClassification
audioSecurity	ivSpkIdFeature	speechToText
clippingDetection	ivSpkIdTrainAudio	speechToTextFilter
clusterSpeech	ivSpkIdDevelAudio	speechToTextTelephony
clusterSpeechTel	langId	
clusterSpeechToTextTel	langIdFeature	

In addition, these tasks now also accept the `StartTime` and `EndTime` parameter to choose a point to start and end in an audio file or stream.

- The new `SpeechToTextTelephony` task now applies music and noise filtering to the transcript, in addition to dial-tone and DTMF detection.
- The `fptdb` resource has been deprecated. You can now use the `fpdb` resource for template audio fingerprinting, as well as standard.
- The `SegmentWav` task has been deprecated. You can use the `ClusterSpeech` task to perform a similar operation.
- Acoustic modeling is now deprecated, in favor of the newer Deep Neural Network (DNN) approaches. The following tasks have been deprecated:
 - `WavToPlh`
 - `AmTrain`
 - `AmTrainFinal`
 - `DataObfuscation`

The following related modules are also deprecated:

- `amadapatadddata`
- `amadaptend`

In addition, the `TrainedAm` task parameter and `TrainedAmDir` configuration parameter are now deprecated.

- The `lbout` module has been deprecated. Information about language boundaries is provided in the output from the `lidout` module. Speech Server language identification boundary mode no longer produces a separate boundary output file.
- The audio fingerprinting tasks have a new `OutputNonResults` parameter to allow you to output periodic non-results options in periods where no matches are detected. This option provides feedback on the progress of the matching process.
- You can now force Speech Server to unload a language pack immediately by setting the new `Force` parameter to `True` in the `UnloadLanguage` action.
- A new base language identification pack and classifier set has been added for broadband language identification, taking advantage of new neural network technology.

NOTE:

If you use user-trained language classifiers, HPE recommends that you retrain your

classifiers with the new base pack. If you do not retrain the classifiers, performance might be adversely affected.

- You can now configure GSS authentication on the ACI and service ports without using ACI encryption. In this mode, all connections to the ports must be authenticated using GSSAPI and the Negotiate HTTP authentication mechanism.

To use GSS authentication, you must set the `GSSServiceName` parameter in the `[Server]` section to the full service name, domain, and Kerberos realm for the service. You can then set the `RequireGSSAuth` parameter in the `[Server]` section to enable GSS authentication on the ACI port, and set `RequireGSSAuth` in the `[Service]` section to enable GSS authentication on the service port.

NOTE:

You cannot configure `RequireGSSAuth` with the `[ACIEncryption]` configuration options. If you attempt to configure both, the server does not start.

This method provides an authentication requirement only. HPE recommends that you use it in conjunction with TLS/SSL to encrypt the authentication data.

- All ACI server ports now support the `Expect: 100-continue` HTTP header. Previously, third-party client applications that used this header (for example, using the cURL utility with the `-F` option to POST form data) could experience increased latency when communicating with the ACI server.

Resolved Issues

This section lists the resolved issues in IDOL Speech Server version 11.5.0.

- In some cases, it was not possible to retrieve results from Speaker ID and Language ID tasks by using the `GetResults` action when the returned labels (speaker or language names) contained spaces.

NOTE:

Speech Server now writes any labels that contain spaces into the CTM results file using the ` ` HTML identifier to replace the space. The space is restored when you use `GetResults` to view the results.

- The task progress time that was reported by `GetStatus` (in the `ProcessingEnd` value) could occasionally regress.
- Speech Server could send warnings when using an audio fingerprinting database that had previously had tracks removed.
- Speech Server could sometimes exit unexpectedly while running DTMF or dial tone identification.
- The `TaskHelp` action could log spurious errors.
- When an authorization role defined `Actions`, `ServiceActions`, or `IndexActions`, and the authorization role `Clients` parameter contained host names, calling the `ShowPermissions` action could result in an interruption of service.

Notes

These notes provide extra information about installing and using IDOL Speech Server.

- Installation on Linux requires the following software:
 - GLIBC_2.3.2
 - GLIBCXX_3.4.20
 - GCC_4.8.0
- If you install IDOL Speech Server 11.5.0 using the IDOL 11.5.0 installer program, you must ensure that you have a Speech Server license key in addition to the standard IDOL Server license key. The IDOL Server license key does not contain licensing information for Speech Server, and Speech Server cannot run using it.
- The Solaris operating system does not support the audio fingerprinting feature in Speech Server.
- The following standard tasks have been deprecated:
 - SpkIdDevel
 - SpkIdDevelFinal
 - SpkIdDevelStream
 - SpkIdDevelWav
 - SpkIdEvalStream
 - SpkIdEvalWav
 - SpkIdFeature
 - SpkIdSetAdd
 - SpkIdSetDelete
 - SpkIdSetEditThresh
 - SpkIdSetInfo
 - SpkIdTmpEditThresh
 - SpkIdTmpInfo
 - SpkIdTrain
 - SpkIdTrainStream
 - SpkIdTrainWav

Use the equivalent iVector tasks instead (for example IvSpkIdDevel). These tasks are still available for existing implementations, but they might be incompatible with new functionality. The tasks might be deleted in future.

- The following standard tasks were deprecated in earlier versions of Speech Server. The documentation for these tasks has now been removed:
 - SidPackage
 - SidTrain
 - SidTrainFinal

- StreamSidOptimize
- StreamSidTrain
- StreamSpeakerId
- WavSidOptimize
- WavSidTrain
- WavSpeakerId

In addition, the documentation for the following associated modules, which were also deprecated, has been removed:

- sidfeature
- sidtrain
- sidoptimizer
- sidpackager
- The following action parameters were deprecated in earlier versions of Speech Server. The documentation for these parameters has now been removed:
 - ClassPrefix
 - Norm
- The following configuration parameters were deprecated in earlier versions of Speech Server. The documentation for these parameters has now been removed:
 - [sidout] module FullInfo
 - [Paths] TasksConfig
 - [Server] CustomLMDir
 - [Server] NestedStatus
 - [Server] TempDir
 - [Server] TrainedAmDir
- Documentation for the following tasks has been removed. These standard tasks were removed from the Speech Server configuration file in a previous version. You can now add punctuation in the relevant speech to text tasks by using the Punctuation parameter.
 - StreamToTextMusicFilterPunct
 - StreamToTextPunct
 - TelWavToTextPunct
 - WavToTextPunct
- The following configuration parameters for setting server action authorization by client IP address have been deprecated:
 - [Server] AdminClients
 - [Server] IndexClients
 - [Server] QueryClients or UserClients
 - [Service] ServiceControlClients
 - [Service] ServiceStatusClients

You can now use the [AuthorizationRoles] configuration section to set up authorization for your servers more flexibly. These configuration parameters are still available for existing

implementations, but they might be incompatible with new functionality. The parameters might be deleted in future.

Documentation

The following documentation was updated for this release.

- *IDOL Speech Server Administration Guide*
- *IDOL Speech Server Reference*