



Hewlett Packard
Enterprise

HPE Speech Server

Software Version: 11.1.0

IDOL Speech Server Release Notes

Document Release Date: June 2016
Software Release Date: June 2016

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 2016 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

HPE Big Data Support provides prompt and accurate support to help you quickly and effectively resolve any issue you may encounter while using HPE Big Data products. Support services include access to the Customer Support Site (CSS) for online answers, expertise-based service by HPE Big Data support engineers, and software maintenance to ensure you have the most up-to-date technology.

To access the Customer Support Site

- go to <https://customers.autonomy.com>

The Customer Support Site includes:

- **Knowledge Base.** An extensive library of end user documentation, FAQs, and technical articles that is easy to navigate and search.
- **Support Cases.** A central location to create, monitor, and manage all your cases that are open with technical support.
- **Downloads.** A location to download or request products and product updates.
- **Requests.** A place to request products to download or product licenses.

To contact HPE Big Data Customer Support by email or phone

- go to <http://www.autonomy.com/work/services/customer-support>

Support

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 updates or to verify that you are using the most recent edition of a document, visit the Knowledge Base on the HPE Big Data Customer Support Site. To do so, go to <https://customers.autonomy.com>, and then click **Knowledge Base**.

The Knowledge Base contains documents in PDF and HTML format as well as collections of related documents in ZIP packages. You can view PDF and HTML documents online or download ZIP packages and open PDF documents to your computer.

Contents

- New in this Release 5
- Resolved Issues 7
- Notes 8
- Documentation 9

New in this Release

This section lists the enhancements to Speech Server version 11.1.0.

- This release includes a set of tasks that you can use to add punctuation to speech-to-text output:

WavToTextPunct	StreamToTextMusicFilterPunct
StreamToTextPunct	PunctuateCtm
TelWavToTextPunct	

You can use the new `NonSentFinalWords` parameter with any of these tasks to specify a list of words in the recognized language that are highly unlikely to end a sentence.

- This release includes a new set of alternative tasks for audio fingerprinting:

AfptMatchWav	AfptAddTrackStream
AfptMatchStream	AfptRemoveTrack
AfptAddTrackWav	AfptDatabaseInfo

These tasks are based on a more robust but less scaleable algorithm, and as such are more tolerant of echos, audio distortion and so on, but are only useable with smaller datasets.

- You can use the `FrameDupl` parameter to allow for greater time efficiency with only a minimal loss of recognition accuracy. You can use this parameter in the following tasks:

WavToText	StreamToTextMusicFilter
StreamToText	WavPhraseSearch
TelWavToText	WavToFMD

- This release includes three new standard tasks that allow greater control over reviewing and editing the thresholds associated with speaker templates:
 - The `SpkIdSetEditThresh` parameter allows you to modify the threshold of a single template in an audio template set file.
 - The `SpkIdTmpEditThresh` parameter allows you to modify the threshold of a single template by specifying the template file.
 - The `SpkIdTmpInfo` parameter allows you to write information on specified template file to a log file.
- You can use the `TimeQuant` configuration parameter in the `afp` module, and the `TimeQ` action parameter in the `afpMatchWav` and `afpMatchStream` tasks, to control the amount of quantization to use.
- You can use the `MinPhoneRate` parameter in language identification tasks to manipulate the threshold at which a segment is considered valid. You can use this parameter in the following tasks:

LangIdSegWav	LangIdBndStream
LangIdCumWav	LangIdSegLif
LangIdBndWav	LangIdCumLif
LangIdSegStream	LangIdBndLif
LangIdCumStream	

In addition, you can also use the `SpeechThresh` and `SilThresh` parameters in the `LangIdFeature` task and all the tasks mentioned above, to manipulate the threshold between speech and non-speech, or silence and non-silence.

- You can specify the name of the generated Language Model Information file by using the `NewLmInfoFile` parameter. You can retrieve the file by using the `GetResults` action with the `Label` parameter set to `lmi`. You can also use the `NewLmInfoFile` parameter in the `[lmbuild]` module to set the output LMI file name. In addition, the `NewDictionary` parameter in the `[lmbuild]` module and `LanguageModelBuild` task is no longer compulsory.
- You can use the new `DropFrames` parameter in the `[filter]` module to discard filtered frames rather than including them in the output.
- The `ShowPermissions` action now shows the rules that define whether a particular origin IP has a particular type of permission. This information is returned only if you send the `ShowPermissions` action from a client IP that is configured in the `AdminClients` configuration parameter.

Resolved Issues

This section lists the resolved issues in Speech Server version 11.1.0.

- The default value for the `DataExt` action parameter and the `DataListExt` configuration parameter has been changed from `.plh` to `.atv`. This affects the `SpkIdDevel` and `SpkIdTrain` tasks, and the `audiotemplatedevel` and `audiotemplatetrain` modules.
- Segmentation of Chinese, Japanese, and Korean now uses the latest Rosetta Language Processing (RLP) libraries, to resolve an issue with loss of service when using older libraries.
- An issue was resolved that could lead to loss of service when outputting multiple speech-to-text lattice files simultaneously across multiple tasks.
- The default values for the `LabPath`, `LabExt`, and `Format` parameters in the `audioTemplateDevel`, `audioTemplateTrain`, and `sidout` modules have been updated. In addition, the `FullInfo` parameter in the `sidout` module has been deprecated.
- In Speaker Identification, the speaker name in non-speech sections is now `NonSpeech_` instead of `Unknown_`.
- LMI files generated by `LanguageModelBuild` tasks are now deleted when you call the `DeleteResults` action, as well as TLM and DCT files.
- An issue was resolved that could lead to errors whilst running multiple speaker ID development tasks in parallel.
- Audio fingerprint matching now gives an error if the supplied database is empty (that is, if it contains no audio tracks), instead of processing the data regardless of the fact that it will never produce any results.
- ENUS normalization now uses the correct format for spelling out months. In addition, `1000` is correctly normalized as *one thousand*, and `2:00 pm` is now correctly spelled as *two pm* instead of *two zero pm*.
- The language model loading process is now more robust in terms of handling network errors.
- The `ShowPermissions` action did not return details for `IndexClients`, `ProxyClients`, `ServiceStatusClients`, and `ServiceControlClients` if these values were not explicitly set in the configuration file.

Notes

These notes provide extra information about installing and using 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 11.1.0 using the IDOL 11.1.0 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.

Documentation

The following documentation was updated for this release.

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