



Private & Confidential

## Intelligent Voice 5.0 Release Note

Date: 15<sup>th</sup> June 2021

Intelligent Voice Limited  
St Clare House  
30-33 Minories  
London EC3N 1DD  
t: +44 20 36 27 26 70  
e: [info@intelligentvoice.com](mailto:info@intelligentvoice.com)  
w: [www.intelligentvoice.com](http://www.intelligentvoice.com)



## INTELLIGENT VOICE DISCLAIMER

© 2021 Intelligent Voice. All rights reserved.

Intelligent Voice believes the information in this publication is accurate as of its publication date. The information is subject to change without notice. The information in this publication is provided “as is”. Intelligent Voice makes no representation or warranties of any kind with respect to the information in this publication and specifically disclaims implied warranties of merchantability or fitness for a particular purpose. Using, copying, and distribution of any Intelligent Voice software described in this publication requires an applicable software license.

THE INFORMATION CONTAINED IN THIS DOCUMENT IS THE PROPRIETARY AND CONFIDENTIAL INFORMATION OF INTELLIGENT VOICE LIMITED. YOU MAY NOT DISCLOSE, PROVIDE OR MAKE AVAILABLE THIS DOCUMENT, OR ANY INFORMATION CONTAINED IN THIS DOCUMENT, TO ANY THIRD PARTY, WITHOUT THE PRIOR WRITTEN CONSENT OF INTELLIGENT VOICE LIMITED.

The information in this document is provided for use with Intelligent Voice products. No licence, express or implied, to any intellectual property associated with this document or such products is granted by this document.

All Intelligent Voice products described in this document are owned by Intelligent Voice Limited (or those companies that have licensed technology to Intelligent Voice) and are protected by patents, trade secrets, copyrights or other industrial property rights.

The Intelligent Voice products described in this document may still be in development. The final form of each product and release date thereof is at the sole and absolute discretion of Intelligent Voice. Your purchase, license and/or use of Intelligent Voice products shall be subject to Intelligent Voice’s then current sales terms and conditions.

## Table of Contents

<b>New Features &amp; Improvements</b> .....	5
Functional .....	5
Improved Automatic Speech Recognition.....	5
Updated Diarization (Biometric Speaker Separation) .....	5
Punctuation .....	5
NVIDIA Triton Inference Server .....	5
Faster waveforms in SmartTranscripts (Preview Feature).....	5
Technical .....	6
Merge of CPU & GPU ASR Workers.....	6
Containerisation of Intelligent Voice Workers .....	6
Worker Monitoring.....	6
Upgrade to Drupal 7.7.5 .....	6
Upgrade to Apache Tomcat 8.5.....	6
Upgrade to Python 3 .....	6
Ubuntu 20.04 (preview support).....	6
Upgrade to ffmpeg 4 .....	6
<b>Bug Fixes</b> .....	7
Multilingual Tagger.....	7
Video Retrieval via API.....	7
Gearman Server / Processing Node Restarts .....	7
NullPointerException for Some Video Formats .....	7
Some non-.wav Audio Files Not Converting .....	7
Support for MariaDB Connector/J 2.7 and later .....	7
Components Not Restarting After Reboot on RedHat .....	8
<b>Known Issues</b> .....	8
Queries ran by schedulers get stuck .....	8
Psrt Mangler concatenates digits into a single unique word .....	8
VRX Certificate Not Verified .....	8
Unencrypted Gearman Connections.....	9
Some Items Created with Base64 Fail Format Conversion .....	9
MySQL Unable to Automatically Restart on RedHat Enterprise Linux.....	9

GPU Contention between nASR and Triton Inferencing .....	9
Triton Server Connection Not Authenticated and Not Encrypted.....	9
Language Models & Language Model Builder Installed with Wrong Permissions.....	10
Redacted Text is Available as a Topic and Alternatives in Exported .html SmartTranscript .....	10
getGroupDetails Times Out For Over 1 Million Items.....	10
Some Video Files Are Not Exported as Video in the .html SmartTranscript .....	10
Waveform Can Fail to Generate On Screen .....	11
Waveform Not Generated for Video When Calling the Get Item webservice .....	11
Online Installation Looks for _python-packages Folder .....	11
SphinxSearch Index Rotation Fails on Empty Groups .....	11
Installer Fails if User selects a Password with @.....	11
Hyphenated Alternatives Do Not Work for Japanese transcripts.....	11
Files with No Voice Detected are Not Marked as Complete.....	12
Docker Components Running As Root .....	12
Docker Security Audit Actions.....	13
Item Method Can Produce Invalid JSON.....	13
Misrepresentation of Background Talk & Over Talk with nASR .....	13
/recordings/ Endpoint Returns HTTP 410 When It Can't Convert a File .....	14
Import Report Search changes From/To date automatically .....	14
Items Without a "from" or "to" Populated Are Not Added to the Search Index.....	14
Import Report Search changes From/To date automatically .....	14
Topics Not Created for Mandarin .....	15
Tagger Errors when Unicode Characters Encountered .....	15
Numification When Active Turns "for" into "4" Throughout.....	15
Import Treating .csv Metadata File as an Item.....	15
Japanese, Russian, Korean, Arabic Tags cannot be seen via jump2web .....	15
Karaoke Highlighting of Index Text Misaligned with Speech for nASR.....	15
Installing Intelligent Voice 5.0 Disconnects Machine From the Network if Subnet 172.17.0.0/16 is used ..	16
Intelligent Voice Cannot Be Installed on a Server With More Than 64 Characters in its Hostname .....	16

## **New Features & Improvements**

### **Functional**

#### **Improved Automatic Speech Recognition**

New, CTC-based Automatic Speech Recognition (nASR) initially available for English, offering improved accuracy over existing IV models.

#### **Updated Diarization (Biometric Speaker Separation)**

The previous i-vector based Diarization system is replaced by an x-vector based system for improved speaker separation accuracy across all languages.

#### **Punctuation**

Further advances in punctuation, now including commas and question marks in English, together with full stops (periods) across multiple languages.

#### **NVIDIA Triton Inference Server**

New support for running AI workloads with the addition of NVIDIA Triton Inference Server (<https://developer.nvidia.com/nvidia-triton-inference-server>).

#### **Faster waveforms in SmartTranscripts (Preview Feature)**

Waveforms in SmartTranscripts are now pre-generated, rather than created on-screen with a dependency on JavaScript. Only supported for audio files, not video and only when installed on Ubuntu 20.04 - itself in preview / beta – and not RedHat Enterprise Linux.

## Technical

### **Merge of CPU & GPU ASR Workers**

There is now a single Automatic Speech Recognition worker for both CPU and GPU based processing.

### **Containerisation of Intelligent Voice Workers**

Core Intelligent Voice Workers are now published as Docker containers on the Intelligent Voice container registry at [ivcr.io](https://ivcr.io):

- VAD – Voice Activity Detection
- ASR – Automatic Speech Recognition
- Diarization – Biometric Speaker Separation
- Cracker – Text File Indexing
- Sphinx – Search Index
- Tagger – Topic Extraction

### **Worker Monitoring**

The Gearman Queue page in the web browser now auto-refreshes every 15 seconds.

### **Upgrade to Drupal 7.7.5**

The Drupal framework used for JumpToWeb has been upgraded to 7.7.5.

### **Upgrade to Apache Tomcat 8.5**

Upgraded to Tomcat Server 8.5.

### **Upgrade to Python 3**

Components previously using Python 2 now Python 3.

### **Ubuntu 20.04 (preview support)**

Preview support for installing Intelligent Voice on Ubuntu 20.04. Note: Official support for Ubuntu 16.04 ended in April 2021 but Intelligent Voice continues to support existing installs to allow customers to upgrade.

### **Upgrade to ffmpeg 4**

ffmpeg has been upgrade to version 4.

## Bug Fixes

### Multilingual Tagger

<b>Items</b>	INT-4396, INT-4115
<b>Description</b>	Tags are now correctly produced for: <ul style="list-style-type: none"> <li>• Belgian Dutch [nl-BE]</li> <li>• Netherlands Dutch [nl-NL]</li> <li>• German [de-DE]</li> <li>• Swiss German [de-CH]</li> <li>• Japanese</li> <li>• Spanish [es-ES]</li> </ul>

### Video Retrieval via API

<b>Items</b>	INT- 4339, INT-4045
<b>Description</b>	Web service for the API call to Get Recording for Item can now return MP4. The itemId endpoint was not returning a waveform for video files.

### Gearman Server / Processing Node Restarts

<b>Items</b>	INT-4050
<b>Description</b>	gearmanstats service restarts automatically when gearman-job-server is restarted.

### NullPointerException for Some Video Formats

<b>Items</b>	INT-4055
<b>Description</b>	Fix to ffmpeg NullPointerException for some video file formats.

### Some non-.wav Audio Files Not Converting

<b>Items</b>	INT-3925
<b>Description</b>	Error message received: <i>System.Net.WebException: The remote server returned an error: (500) Internal Server Error. at System.Net.HttpWebRequest.GetResponse() at IntelligentVoice.ServiceWrapper.HttpRequestHandler.MakeHttpRequestDownloadContent(HttpWebRequest httpRequest)</i>

### Support for MariaDB Connector/J 2.7 and later

<b>Items</b>	INT-3111, INT-4074, INT-4087, INT-4384
<b>Description</b>	Changes to support latest versions of the database driver MariaDB Connector/J version 2.7.2. This update brings performance and security improvements

### Components Not Restarting After Reboot on RedHat

<b>Items</b>	INT-4408, INT-4471
<b>Description</b>	MariaDB not starting after reboot. ASRWorker could not autostart after VM restart.

## Known Issues

### Queries ran by schedulers get stuck

<b>Items</b>	INT-3534
<b>Description</b>	<p>Queries run by a schedule may appear to get stuck as the number of items increase in the system, as the scheduler appears not to be running.</p> <p>The issue was caused by the SQL queries not having a LIMIT. This mostly affects systems with &gt; 1 million items</p>
<b>Workaround</b>	<p>Increase database setting <code>innodb_buffer_pool</code> <i>or</i></p> <p>Review all the queries ran by the schedulers and add a LIMIT, e.g. "LIMIT 1000" . For the queries with UNION, add LIMIT to each one, wrapping them into ( ) brackets to avoid SQL syntax error.</p>

### Psrt Mangler concatenates digits into a single unique word

<b>Items</b>	INT-3606, INT-3703
<b>Description</b>	<p>For example, the statement 'one eighty' would generate both ONE and EIGHTY in the lattice of alternative words, but only "one" would be present in the srt file.</p> <p>The words 'teen' and 'teens' can also be lost between the lattice of alternative words and the srt file when the Psrt Mangler is used.</p>
<b>Workaround</b>	None.

### VRX Certificate Not Verified

<b>Items</b>	INT-2124
<b>Description</b>	The JumpToWeb application does not verify the SSL certificate presented by the vrx application
<b>Workaround</b>	Install JumpToWeb on the same server as the application server, and set the server name to resolve to 127.0.0.1 (this is the recommended setup anyway)



### Unencrypted Gearman Connections

<b>Items</b>	INT-2430, INT-2465
<b>Description</b>	Connections between Gearman clients and the server are not encrypted. The Gearman process parameters include the mysql password.
<b>Workaround</b>	Single-server installs can change Gearman to only accept connections on localhost. Multi-server installs can create a SSH tunnel – contact support for assistance

### Some Items Created with Base64 Fail Format Conversion

<b>Items</b>	INT-4278, INT-4549
<b>Description</b>	If an item is created with invalid base64, this creates an error for the item in the j2ws.log, as the base64 is not pre-validated.
<b>Workaround</b>	None.

### MySQL Unable to Automatically Restart on RedHat Enterprise Linux

<b>Items</b>	INT-3714
<b>Description</b>	Sometimes MySQL is unable to restart after a reboot on RHEL.
<b>Workaround</b>	Manually start MySQL and restart tomcat and gearman by executing the commands: <pre>sudo systemctl enable mariadb sudo systemctl start mariadb sudo systemctl restart tomcat8 gearman-job-server</pre>

### GPU Contention between nASR and Triton Inferencing

<b>Items</b>	INT-4342, INT-4338
<b>Description</b>	Contention between the new, CTC-based alternative model Automatic Speech Recognition (nASR) and Triton Inference Server if installed on the same machine, and so sharing GPU. Currently Triton is using all GPUs in its service file.
<b>Workaround</b>	Use Triton in CPU mode. If you have multiple GPUs you can also temporarily change the actual container as follows (changes will be lost when the container restarts: Create a lock on GPU 0 for ASR as follows: <ol style="list-style-type: none"> <li>1. Change ASR to not use GPU 0             <ol style="list-style-type: none"> <li>a. <code>docker exec -it ASRWorker bash</code></li> <li>b. <code>cd /opt/IV/lock</code></li> <li>c. <code>touch gpu0</code></li> </ol> </li> <li>2. change triton to only use GPU 0 (fill instructions here)</li> </ol>

### Triton Server Connection Not Authenticated and Not Encrypted

<b>Items</b>	INT-4575
--------------	----------

<b>Description</b>	The Automatic Speech Recognition Container makes an unauthenticated, unencrypted connection to the Triton server.
<b>Workaround</b>	Install Triton on the same server as ASR, and set the Triton server address to 'localhost' in the installer config

### Language Models & Language Model Builder Installed with Wrong Permissions

<b>Items</b>	INT-3902
<b>Description</b>	World-writable files can be modified and/or executed by other local users. Language Models could be amended other than by the Intelligent Voice Language Model Builder.
<b>Workaround</b>	Change permissions on the folder for the Language Models /opt/models/* as follows: chown tomcat8:tomcat8 chmod -R 770 /opt/models/* NOTE this must be repeated every time you build or install a new model when the processing node fetches a new model. WARNING for autoscaling images this may not be practical – contact support for more assistance with autoscaling

### Redacted Text is Available as a Topic and Alternatives in Exported .html SmartTranscript

<b>Items</b>	INT-2661
<b>Description</b>	Text redacted in the aligned index text for the player remains unredacted in any Tag containing that text for the portion of audio. The alternative words text for the redacted audio also remains unredacted.
<b>Workaround</b>	Avoid utilising the .html export format for workflows involving redactions. Utilise the combination of audio file and index text (.wav & .txt) instead.

### getGroupDetails Times Out For Over 1 Million Items

<b>Items</b>	INT-2059
<b>Description</b>	Get Group Details can timeout for large groups of over a million items. This can prevent the main JumpToWeb screen loading
<b>Workaround</b>	Use multiple groups to keep item count under 1 million per group For the API call: Retrieve items by a particular date & time range or Item Id range, instead of all items for the Group. There is no workaround for JumpToWeb screen, but individual transcripts can still be accessed by ID

### Some Video Files Are Not Exported as Video in the .html SmartTranscript

<b>Items</b>	INT-3911
<b>Description</b>	Some video file formats are processed as if audio, resulting in the video element not appearing in the exported .html SmartTranscript.

<b>Workaround</b>	For batch processing, pre-convert the video to video to MP4 container, H264 video codec, AAC audio codec, prior to importing it into Intelligent Voice for processing.
-------------------	--

### Waveform Can Fail to Generate On Screen

<b>Items</b>	INT-3897
<b>Description</b>	For audio on Red Hat Enterprise Linux, and video in either RHEL or Ubuntu, the waveform can fail to generate, due to its dependency on JavaScript. This will occur on all files if JavaScript has been disabled in the browser, and can occur on longer files even where JS is enabled.
<b>Workaround</b>	If occurring on all files, allowing JavaScript to execute in the browser will resolve for shorter files.

### Waveform Not Generated for Video When Calling the Get Item webservice

<b>Items</b>	INT-3904
<b>Description</b>	Get Item does not generate a Waveform for video.
<b>Workaround</b>	None.

### Online Installation Looks for `_python-packages` Folder

<b>Items</b>	INT-4024
<b>Description</b>	This is a security issue as the packages are out of date, with security patches not applied. The majority of affected code has been moved into containers.
<b>Workaround</b>	Risk mitigation: validate untrusted inputs before inserting into IV, reduce exposure to untrusted inputs

### SphinxSearch Index Rotation Fails on Empty Groups

<b>Items</b>	INT-3928
<b>Description</b>	Sphinxsearch index rotation stops if it encounters an empty group, meaning higher number groups never get rotated.
<b>Workaround</b>	Populate the empty group with a short, silent file.

### Installer Fails if User selects a Password with @

<b>Items</b>	INT-3749
<b>Description</b>	Installer Fails if User selects a Password with @.
<b>Workaround</b>	Use a different special character for the password instead.

### Hyphenated Alternatives Do Not Work for Japanese transcripts

<b>Items</b>	INT-3749
<b>Description</b>	In the lattice of hyphenated alternatives, all characters are replaced with 'e'.

<b>Workaround</b>	<p>Restricted searching to the displayed index text for Japanese, rather than the lattice.</p> <p>Export only the displayed index text for Japanese, and not the alternatives lattice text.</p>
-------------------	---

### Files with No Voice Detected are Not Marked as Complete

<b>Items</b>	INT-4067, INT-4179
<b>Description</b>	Files with no voice detected, such as silent files or files consisting only of noise, are not marked as complete due to an error in the apache log.
<b>Workaround</b>	None.

### Docker Components Running As Root

<b>Items</b>	INT-4386
<b>Description</b>	A number of Docker components, e.g. Tagger, are running as root, instead of under their dedicated system accounts.
<b>Workaround</b>	None.

### Docker Security Audit Actions

<b>Items</b>	INT-4058
<b>Description</b>	<p>Actions arising from:</p> <ol style="list-style-type: none"> <li>1. The Center for Internet Security's step-by-step checklist to secure Docker <a href="https://www.cisecurity.org/benchmark/docker/">https://www.cisecurity.org/benchmark/docker/</a></li> <li>2. Docker's own "Docker Bench for Security" <a href="https://github.com/docker/docker-bench-security">https://github.com/docker/docker-bench-security</a></li> </ol> <p>Actions arising include:</p> <ol style="list-style-type: none"> <li>1. Ensure that the PIDs cgroup limit is used</li> <li>2. Ensure that container health is checked at runtime</li> <li>3. Ensure that the container is restricted from acquiring additional privileges</li> <li>4. Ensure that the container's root filesystem is mounted as read only</li> <li>5. Ensure that CPU priority is set appropriately on containers</li> <li>6. Ensure that the memory usage for containers is limited</li> <li>7. Ensure that the host's network namespace is not share</li> <li>8. Ensure network traffic is restricted between containers</li> <li>9. Ensure centralized and remote logging is configured</li> <li>10. Ensure that a user for the container has been created</li> <li>11. Ensure live restore is enabled</li> <li>12. Ensure Content trust for Docker is Enable</li> <li>13. Ensure that HEALTHCHECK instructions have been added to container images</li> </ol>
<b>Workaround</b>	None.

### Item Method Can Produce Invalid JSON

<b>Items</b>	INT-4044
<b>Description</b>	Item Method Can Produce Invalid JSON.
<b>Workaround</b>	None.

### Misrepresentation of Background Talk & Over Talk with nASR

<b>Items</b>	INT-4230, INT-4238
<b>Description</b>	Where Voice Activity Detection thinks speech may be present, nASR may insert a series of iiiii's or aaa's or er's for background talk or / and overtalk.
<b>Workaround</b>	As a manual or scripted post-export / production process strip out: <ul style="list-style-type: none"> <li>• 2 or more 'ii' or 'Ii' together</li> <li>• Any standalone 'er'</li> <li>• 2 or more 'er' together</li> <li>• 2 or more 'aa' or 'AA' together</li> </ul>

### [/recordings/ Endpoint Returns HTTP 410 When It Can't Convert a File](#)

<b>Items</b>	INT-4346
<b>Description</b>	/recordings/ endpoint returns a HTTP 410 Gone when it can't convert a file, for example if a non-auto-scaled app server runs out of disk space. When the Voice Activity Detector gets a HTTP 410 it discards the job rather than return an error.
<b>Workaround</b>	None.

### [Import Report Search changes From/To date automatically](#)

<b>Items</b>	INT-3873
<b>Description</b>	When an import report search is run, the first two fields to fill in are the From and To dates. Once you run the report, these automatically change to one day before whatever you'd set them to. The report runs correctly but if you then rerun it, you're actually running across a different date range to that originally entered.
<b>Workaround</b>	To re-run the report, for example to monitor live progress, you have to re-enter the date range each time, instead of being able to just hit search again.

### [Items Without a "from" or "to" Populated Are Not Added to the Search Index](#)

<b>Items</b>	INT-3927
<b>Description</b>	Items without a "from" or "to" populated are not added to the search index.
<b>Workaround</b>	Edit sphinx.config: <pre> &lt;      LEFT JOIN  obsilon.`{\$i}_mail_id_link` mil ON (mil.id = m.id))\ &lt;      LEFT JOIN  obsilon.`{\$i}_person_email_link` pel ON (mil.email_from = pel.id_email)"; --- &gt;      INNER JOIN  obsilon.`{\$i}_mail_id_link` mil ON (mil.id = m.id))\ &gt;      INNER JOIN  obsilon.`{\$i}_person_email_link` pel ON (mil.email from = pel.id_email)";           </pre>

### [Import Report Search changes From/To date automatically](#)

<b>Items</b>	INT-3873
<b>Description</b>	Search endpoint returns inconsistent result after delete items. A search after items have been deleted returns 'Results 0 to 0 items of n results' with warning messages about non-numeric values and undefined offset.
<b>Workaround</b>	None.

### Topics Not Created for Mandarin

<b>Items</b>	INT-4390
<b>Description</b>	The tagger worker doesn't produce tags for Mandarin.
<b>Workaround</b>	None.

### Tagger Errors when Unicode Characters Encountered

<b>Items</b>	INT-3181
<b>Description</b>	The tagger work throws an error when Unicode characters are encountered.
<b>Workaround</b>	None.

### Numification When Active Turns "for" into "4" Throughout

<b>Items</b>	INT-4206
<b>Description</b>	Rather than only turn "for" into "4" when there's a numeric context, when Numification is active it turns all instances of "for" into "4"
<b>Workaround</b>	None.

### Import Treating .csv Metadata File as an Item

<b>Items</b>	INT-4200
<b>Description</b>	When importing by using CSV template file, iv presents "item error" on the import report page although all the audio items are imported successfully.
<b>Workaround</b>	API callers can inspect the response to determine if the error is only affecting the .csv file, and treat it as success.

### Japanese, Russian, Korean, Arabic Tags cannot be seen via jump2web

<b>Items</b>	INT-3828
<b>Description</b>	Opening item details via jump2web only shows a single tag, containing numbers, when item was transcribed in Japanese, Russian, Korean or Arabic.
<b>Workaround</b>	Export a .html SmartTranscript to review the item.

### Karaoke Highlighting of Index Text Misaligned with Speech for nASR

<b>Items</b>	INT-4240
<b>Description</b>	In a .html SmartTranscript created using nASR, the karaoke may highlight a different part of the index to where people are speaking.
<b>Workaround</b>	None.

### Installing Intelligent Voice 5.0 Disconnects Machine From the Network if Subnet 172.17.0.0/16 is used

<b>Items</b>	INT-4380
<b>Description</b>	Due to default Docker bridge network behaviour, where on installation it will create a bridged interface docker0 with a 172.17.0.0/16 subnet for container networking.
<b>Workaround</b>	If possible, change the machine to use a different subnet.

### Intelligent Voice Cannot Be Installed on a Server With More Than 64 Characters in its Hostname

<b>Items</b>	INT-4399
<b>Description</b>	<p>Intelligent Voice generates self-signed certificates as a required part of the installation of app server, database, sphinxsearch, JumpToWeb or Tagger.</p> <p>The self-signed certificate uses the output of the command "hostname -f" as the certificate Common Name and Subject Alternative Name. As part of the ASN.1 standard, a Common Name is limited to 64 chars.</p>
<b>Workaround</b>	Assign the server a hostname of less than 64 chars if possible. Otherwise disable the certificate generation and issue via another means a certificate with less than 64 chars in the Common Name (the Subject Alternative Name can still have the desired hostname over 64 chars allowing the IV install config hostname to have over 64 chars)