



# ISERMO

Speech › Text › Translate

## Operator Guide

Everyday reference for running live captioning and translation

Release 2026.10 · Document revision E · Applies to all ISERMO captioning appliances



80 Little Falls Road • Fairfield, NJ 07004  
800-631-0868 • Sales@HamiltonBuhl.com

[www.HamiltonBuhl.com](http://www.HamiltonBuhl.com)

## Contents

Right-click here and choose “Update Field” to build the table of contents.

### Getting Started

- System Overview
- Logging In
- User Roles

### Daily Operation

- Live Tab
- Sessions
- Mics Tab
- Hotkeys

### Content

- Presets
- Vocabulary
- Glossary
- Archive & Export
- Meeting Minutes & Knowledge Base

### Audience & Displays

- Audience View & Listen
- TV Screens (TV and TV2)
- Wi-Fi QR
- TV Idle Pages
- Screen Share

### Tools

- Demo Mode
- Upload Audio
- Bluetooth Mics
- Health Tabs
- Safe Shutdown

### Help

- Troubleshooting
- Glossary of Terms

Release 2026.10 · Document revision E · Applies to all ISERMO captioning appliances

## System Overview

The ISERMO appliance turns live speech into on-screen captions and translations within a couple of seconds. Everything runs on the appliance itself: audio, speech recognition, translation, spoken audio and meeting minutes are all processed on the unit, and nothing is sent to the cloud.

### Tip: How it works

🔊 Microphones → Beta voice detection → Alpha / Delta3 speech recognition → Omega translation → 📺 TV screens, 📞 audience phones and 🗣️ Lambda spoken audio

| Part                      | What it does                                                                                                      |
|---------------------------|-------------------------------------------------------------------------------------------------------------------|
| Operator Console          | This web interface. It controls sessions, mics, languages, displays and settings.                                 |
| TV Display (/tv and /tv2) | Large-format caption screens for the room. Two independent screens are supported.                                 |
| Audience View (/audience) | A phone-friendly page that attendees open by scanning a QR code. It can also read captions aloud.                 |
| Mobile Remote (/mobile)   | A compact remote control for a phone or tablet: live view, mics, session, demo and health.                        |
| Archive                   | A searchable record of every completed session, with exports, Meeting Minutes and a Knowledge Base.               |
| Alpha speech engine       | The fast recognition engine for 25 European languages.                                                            |
| Delta3 speech engine      | A global recognition engine for 30 languages. It includes automatic language detection (the <b>Auto</b> setting). |
| Omega translation         | Real-time translation of every caption into the session's second language.                                        |
| Lambda spoken audio       | Natural voices that read the captions aloud on audience phones ( <b>Listen</b> ).                                 |
| Kappa speaker ID          | Optional. Names the person speaking from their voice when several people share one mic.                           |
| Theta minutes             | Drafts formatted meeting minutes from a finished session, on the appliance.                                       |

### Note: Interface languages

The console, the TV screens, the audience page and this guide are available in English, Español, Português (Brasil), Français, Deutsch, Italiano, 日本語, 한국어 and 简体中文. Use the language selector at the top of each page.

## Logging In

Open the operator address from any modern browser on your network. It is usually `https://<unit-name>.isermo.com/operator/` over the internet, or `https://<unit-name>.isermo.local/operator/` on the unit's own Wi-Fi.

1. Enter your username and password.
2. Click **Sign in**. You land on the Live tab.
3. You stay signed in for about three hours of activity. The console quietly renews your sign-in in the background, so it will not log you out mid-event.

### Caution: Help pages need a sign-in

This guide is only available to signed-in users. If you open a help link while signed out, you are sent to the sign-in page and then brought back to the guide.

### Tip: Blank page?

Press **Ctrl+Shift+R** (Cmd+Shift+R on a Mac) to reload. Chrome, Edge, Safari and Firefox are all supported.

## User Roles

| Role     | Typical user                  | What they can do                                                                                                                        |
|----------|-------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------|
| viewer   | Observer                      | Read-only console access.                                                                                                               |
| operator | Event technician or volunteer | Runs events: sessions, mics, presets, vocabulary, glossary, TV controls, demo, upload audio, plus any extra tabs an admin grants.       |
| admin    | Customer administrator        | Everything an operator can do, plus users, languages, themes, network, Advanced settings, backups, signage, API keys and Safe Shutdown. |

Tabs you are not allowed to use are hidden. The appliance also enforces the same rules on the server, so a hidden tab cannot be reached by typing its address.

## Live Tab

The Live tab is your main view during an event. It mirrors what the TV screens show.

### Caption display

Each caption shows the spoken language on top and its translation below. Captions that the engine is less sure about are shown dimmed. When Speaker ID is on, each caption carries the speaker's name (or "Speaker 1", "Speaker 2"... for voices that are not enrolled).

### Status indicators

| Status                                                                                         | Meaning                                                                               |
|------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------|
|  LIVE         | A session is running and captions are flowing.                                        |
|  READY / idle | A session is running but nobody has spoken recently.                                  |
|  No session   | Nothing is running. Start one on the Session tab.                                     |
|  HOLD         | Captions are held (H key). They are still processed but are not shown on the TVs yet. |
|  Rehearsal    | A test run. Captions show everywhere, but nothing is recorded.                        |

### Editing on the fly

Use **E** to correct the most recent caption and **X** to hide it from the screens. A hidden caption is still kept in the archive, and an edited caption replaces the original on every screen.

## Sessions

A session is one continuous captioning event. Everything said during a session is kept together in the Archive.

### Pre-event check

Before going live, open **Session** → **Pre-event check** and click **Run**. It checks the database, the processing pipeline, the speech, translation and voice engines, the licence, and mic levels. Every line should be green. A yellow line is a warning you can usually proceed with. A red line must be fixed first.

### Starting a session

1. Open the **Session** tab.
2. Enter a session name (for example "Council Meeting – 14 Oct") and, if you like, a room.
3. Optionally change the primary and secondary language for this session only.
4. Tick **Rehearsal** if this is only a test.
5. Click ►**Start Session**. Captions appear within a second or two of speech.

### Ending a session

Click  **End Session**. The transcript and the recorded audio are saved to the Archive automatically. Always end a session with the button: closing the browser does not end it.

### Rehearsal mode

Captions stream to every screen exactly as in a real event, but nothing is written to the Archive. Use it for sound checks, volunteer training and trying new mic layouts. Starting or ending a rehearsal also clears that session's vocabulary.

## Mics Tab

Each row is one audio channel. Changes save immediately and apply live. You never need to restart anything.

### Note: Licensed channels

The banner at the top shows how many channels this unit is licensed for (for example "Licensed: 4 channel(s)"). Channels above that number are hidden. Contact ISERMO to add channels.

| Column / control                                                                    | What it does                                                                                                                                             |
|-------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------|
| #                                                                                   | Hardware channel number (fixed).                                                                                                                         |
| Name                                                                                | Internal label for the channel.                                                                                                                          |
| Speaker                                                                             | The name shown on captions, TVs and in the archive (for example "Mayor Lopez"). It also becomes a spelling hint for the speech engine.                   |
| Source                                                                              | The physical input feeding this channel. It updates by itself when a mic is connected or paired.                                                         |
| Role                                                                                | speaker (a person, the default), translator (a live interpreter feed) or music (non-speech). Travels with each caption as a label.                       |
| Level                                                                               | Live input meter. Speak into the mic and confirm it moves.                                                                                               |
| Gain dB                                                                             | Software gain applied inside the appliance. Raise it for a quiet source, lower it if the meter pins. It does not change the mixing console's own preamp. |
| Mute                                                                                | Stops this channel from being captioned.                                                                                                                 |
| Source lang                                                                         | The language the speaker uses: a fixed language or <b>Auto</b> . See below.                                                                              |
| Translate to                                                                        | The translation target for this mic. "Session default" translates to the other session language.                                                         |
| TV / Aud                                                                            | Show this channel on the TV screens and/or the audience page.                                                                                            |
| Mask                                                                                | Apply the profanity filter to this channel.                                                                                                              |
| Test                                                                                | Prompts "Speak into mic N" so you can identify a channel by watching its meter.                                                                          |
|  | Per-mic vocabulary: names and terms that only this speaker uses.                                                                                         |

### Source language: fixed or Auto

| Setting | Behaviour | Engine used |
|---------|-----------|-------------|
|---------|-----------|-------------|

| Setting                          | Behaviour                                                                                                                                       | Engine used                                           |
|----------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------|-------------------------------------------------------|
| A fixed language (en, es, de...) | The engine expects exactly that language. It is the fastest and most accurate option.                                                           | Alpha for its 25 European languages, otherwise Delta3 |
| Auto (Global – Delta3)           | The language is detected automatically for every sentence. Use it when a speaker switches languages or you do not know the language in advance. | Delta3                                                |

**Tip: Choosing**

If a speaker always uses one language, pick that language. If a speaker switches languages, or you do not know the language in advance, use **Auto**. Your administrator can limit which languages Auto is allowed to report (Advanced → Auto Detect), which stops stray noises from being captioned as a random language.

Use the bulk buttons at the bottom to set every mic at once: **all to Auto**, **all to the primary language** or **all to the secondary language**. Languages that the current engine setting cannot handle are greyed out.

**Pinning a source to a channel**

The appliance finds audio sources automatically and assigns them to channels. To fix a particular source to a particular channel, pin it in the device table. Pins survive restarts and take effect within about five seconds. Clear the pin to go back to automatic assignment.

## Hotkeys

Hotkeys work when the Live tab has focus.

| Key   | Action                                                                                  |
|-------|-----------------------------------------------------------------------------------------|
| H     | Hold / release. Captions keep processing, and releasing shows everything that was held. |
| E     | Edit the last caption.                                                                  |
| X     | Hide the last caption from the screens (it stays in the archive).                       |
| O     | Switch between one-line and two-line layout.                                            |
| + / - | Make the TV caption text larger or smaller.                                             |
| ?     | Show the hotkey overlay.                                                                |

**Tip: Hold**

Use **H** during music, prayer or long pauses to keep the screens clean. Press **H** again when speech resumes.

## Presets

A preset stores a complete mic layout: names, speakers, languages, gains and display ticks.

1. Set up the Mics tab exactly as you want it.
2. Open **Presets**, type a name (for example "Tuesday Council") and click **Save current config**.
3. To reuse it later, click **Load** next to the preset. Every mic updates immediately.

## Vocabulary

The Vocabulary tab lists words that help the speech engines spell names and special terms correctly. Speaker names from the Mics tab are added automatically. Add anything else the room will say (people, places, product names, acronyms) before the event starts.

**Tip: Best results**

Use full names ("Councillor Amara Okafor", not "Amara"). Put one term per line. The vocabulary belongs to the session and is cleared when a rehearsal starts or ends.

## Glossary

The Glossary controls translation. It is kept between sessions. Use it to:

- force a fixed translation for a word or phrase ("Town Hall" → "Ayuntamiento"),
- keep a brand or proper name untranslated (do-not-translate),
- correct recurring wording.

**Bulk import** accepts a CSV or Excel (XLSX) file, or a pasted list. The recognised columns are `term` (or `name` / `street_name`), `category`, `lang` and `aliases`. Terms can be filtered, enabled, disabled or deleted in bulk, and the whole glossary can be exported to CSV.

## Archive & Export

Every saved session appears on the **Archive** page. Operators need an administrator to switch on archive access for their account. Choose a session to read its transcript, play or download the recording, and export.

| Format                                                                                         | Best for                                 |
|------------------------------------------------------------------------------------------------|------------------------------------------|
|  PDF          | A clean, printable transcript.           |
|  Word (.docx) | An editable transcript.                  |
|  SRT          | Subtitles for video editors and players. |
| JSON                                                                                           | Structured data for integrations.        |
|  TTML / IMSC1 | Broadcast-standard timed text.           |
| ZIP                                                                                            | Everything in one download.              |

Deleting a session permanently removes its transcript, captions, minutes and audio.

## Meeting Minutes & Knowledge Base

### Meeting Minutes (Theta)

Turn a finished session into formatted minutes automatically. The on-board Theta writer reads the transcript and drafts the minutes into your template. The transcript never leaves the appliance.

1. In the Archive, open the session and click  **Generate Meeting Minutes**. This takes a few minutes.
2. Review and edit the draft, then click  **Save draft**.
3. Click  **Finalize** to lock the document. Use  **Reopen** if you need to change it again.
4. Download it as **DOCX** (editable) or **PDF**.

#### Caution: Timing

Generate minutes after the event rather than during a live session, so the live captions keep all of the appliance's processing power.

### Report templates

A built-in template is always available. Admins can download the sample, add the organisation's headings and logo while keeping the `{{...}}` placeholders, and upload it as the house template.

### Knowledge Base

Searches every meeting at once: the original text, the translation and the speaker. The best matches come first, and each match jumps straight to that moment in the session.

## Audience View & Listen

Attendees scan the QR code on the TV (or join the unit's Wi-Fi) and open the audience page on their phone. No app or login is needed.

- The layout matches the TV: the spoken language on top and the translation below, each with its language name. When the mics are on Auto, the language names appear as soon as someone speaks.
- **View:** "Primary + Translation" or "Translation only".
- **Font size**, dark/light theme, high contrast, reduced motion, auto-scroll and "Jump to live" are in the accessibility menu. The header tucks away while reading and returns with a swipe down.
- The page itself can be shown in any of the nine interface languages.

### Listen (Lambda spoken audio)

The  **Listen** button reads captions aloud in a natural voice, so attendees can follow with earphones.

- The default voice, "**Same as translated**", follows the translated language at the bottom of the screen. If that language changes, the voice follows it.
- A listener can pick another voice language from the list. Only languages that actually have a voice are offered. Each listener's choice is their own and does not change the screen text.
- If a language has no voice at that moment, a "Voice unavailable" badge appears next to the button.
- To keep the unit responsive, only a few extra voice languages can be active at once. When the limit is reached, a listener sees "Too many voice languages are in use right now".

Administrators can show or hide the Listen button for everyone (Advanced → Settings → Audience Spoken Audio). Open pages update within about a minute.

## TV Screens (TV and TV2)

Two independent screens are supported: **/tv** and **/tv2**. Each has its own toolbar in the operator console, and every change applies instantly.

| Control                | What it does                                                                                                                                                                              |
|------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Orientation            | Landscape or portrait. Portrait can rotate +90° (clockwise) or –90° (counter-clockwise) to suit how the panel is mounted.                                                                 |
| Caption / speaker size | Text sizes, from extra-small up to extra-large.                                                                                                                                           |
| Wi-Fi QR               | Show or hide the "Join Wi-Fi" QR code on this screen.                                                                                                                                     |
| Audience QR            | Show or hide the "Scan for captions" QR code on this screen.                                                                                                                              |
| Disclaimer             | An optional line of text centred at the bottom (for example "Captions are machine-generated and may contain errors"). Each screen has its own text, and the TV text can be copied to TV2. |
| Controls overlay       | Show or hide on-screen hints.                                                                                                                                                             |

When no session is running, the TV shows its idle page (see TV Idle Pages) or a clean READY screen.

## Wi-Fi QR

Shows a Wi-Fi join code on the TV so attendees can reach the unit's own network. This works even without internet. When the unit's hotspot is on, the QR is matched to it automatically ("✓Matched to the ISERMO hotspot"). If the hotspot is off, type the venue's guest Wi-Fi details instead. **Use ISERMO hotspot settings** re-syncs the QR with the hotspot in one click.

## TV Idle Pages

Upload a full-screen page (welcome slide, schedule or sponsor logo) that the TVs show between sessions. There are four slots: TV landscape, TV portrait, TV2 landscape and TV2 portrait. Upload a single self-contained HTML file with its images embedded. TVs update immediately. Use **Clear** to go back to the plain screen, or switch idle pages off to send the TVs straight to READY.

## Screen Share

Mirror a laptop screen onto the TV for classrooms and training. The shared screen fills the top three-quarters and live captions stay in the bottom quarter.

1. Open **Screen Share** and click ►**Start sharing**. Choose a screen, window or tab.
2. Tick **Smoother** only for video. The default is tuned for slides and text.
3. Click **Stop sharing** to return to captions.

Screen sharing needs a recent desktop version of Chrome, Edge or Firefox.

## Demo Mode

Plays a recorded clip through the whole system so every screen lights up just as in a real session. It is useful for demonstrations, training and stage checks. Pick a clip from the demo library (or upload an MP3/WAV), choose the target-language rotation, and click ►**Start**. While a demo is playing, the room mics are automatically ignored, so the demo audio coming out of the speakers is not captioned twice. Click ◻ **Stop** when you are done.

## Upload Audio

Transcribe and translate recordings made elsewhere, such as interviews or past meetings.

1. Set the default target language and, optionally, a room and title prefix.
2. Drop up to 10 files (up to 200 MB each; wav, mp3, m4a, flac, ogg...).
3. Click **Start batch**. You can leave the page; processing continues in the background.
4. When a file shows **done**, click **Open in Archive**.

Each file's language is detected automatically. Non-English files are translated to English, and English files to the target language you chose. Batch work runs in the background at low priority and does not disturb a live event.

## Bluetooth Mics

Pair a Bluetooth microphone or headset: put it in pairing mode, click **Scan & Pair Mic**, then **Pair** and **Connect**. It appears as a source on the Mics tab. For multi-mic events, wired or networked inputs remain the most reliable.

## Health Tabs

These are read-only dashboards. Looking at them never disturbs a session.

| Tab                   | What it shows                                                                                                                                                                                |
|-----------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Alpha · Delta3 Health | The speech engines: quality per mic, repeated-phrase filtering, blocked phrases and recent drops. Admins can manage the blocked-phrase list here.                                            |
| VAD Monitor           | Beta voice detection per channel: is each mic being heard?                                                                                                                                   |
| Linux DSP             | Per-channel clean-up (high-pass → gate → EQ → compressor → gain). Bypass it when a mixing console already processes the audio.                                                               |
| Linux Health          | The appliance's vital signs: services, CPU, memory, disks, AI accelerator, voice engine, network and recent warnings. Optional features that are switched off show as grey <b>disabled</b> . |

## Safe Shutdown

Administrators can power the appliance down safely from the console (Screen Share tab) or the mobile remote. Every service stops in order and the disks are flushed. Never pull the power during a session. After a safe shutdown, press the power button to start the unit again.

## Troubleshooting

### No captions

1. Is a session running? (The Session tab shows LIVE.)
2. Does the mic's Level meter move? Is it muted?
3. Run the Pre-event check and look for red lines.
4. Check the VAD Monitor to see whether the mic is being heard.

### Captions in the wrong language

Set the mic to a fixed language, or ask your admin to limit Auto Detect to the languages actually used in the room.

### Captions late (more than 5 seconds)

One to two seconds is normal. Ask your admin to check Linux Health. Avoid generating minutes during a live event.

### Translation missing

Check that the session has a secondary language and that the mic's "Translate to" is not blank.

### The same sentence appears twice

Two mics are probably picking up the same voice, or the demo audio is playing through the speakers while a live mic is open. Mute the unused mic, or lower its gain.

### Listen has no sound

Tap the page once (phones need a tap before they allow audio), check the phone's volume, and make sure the Listen button has not been hidden by the admin.

### Public link not working

Advanced → Public Access should show the tunnel as active. If not, contact your administrator.

## Glossary of Terms

| Term       | Meaning                                                                            |
|------------|------------------------------------------------------------------------------------|
| ASR        | Automatic speech recognition: speech to text (Alpha and Delta3 engines).           |
| Alpha      | ISERMO's fast speech engine for 25 European languages.                             |
| Delta3     | ISERMO's global speech engine for 30 languages, with automatic language detection. |
| Omega      | ISERMO's translation engine.                                                       |
| Lambda     | ISERMO's spoken-audio (text-to-speech) engine behind Listen.                       |
| Kappa      | ISERMO's speaker identification engine.                                            |
| Theta      | ISERMO's on-board meeting-minutes writer.                                          |
| Beta       | ISERMO's voice-activity detector, which separates speech from silence.             |
| Session    | One captioning event with its own transcript.                                      |
| Rehearsal  | A session that shows captions but records nothing.                                 |
| Hold       | Pause the screens without losing captions (H).                                     |
| Vocabulary | Spelling hints for names and terms (per session).                                  |
| Glossary   | Translation rules and do-not-translate terms (permanent).                          |
| Confidence | How sure the engine is; low-confidence text is dimmed.                             |



80 Little Falls Road • Fairfield, NJ 07004  
800-631-0868 • Sales@HamiltonBuhl.com

**[www.HamiltonBuhl.com](http://www.HamiltonBuhl.com)**