Your content, copyright and AI

What happens to the media you upload: who owns it, where it goes, who can see it, how long we keep it, and whether it is ever used to train AI.

Last updated

29 September 2026

You keep every right

Your media and everything you make from it belong to you. We claim no ownership and no licence beyond running the service.

No AI training without your agreement

We only train or improve AI models on your content if your organisation agrees to it in writing. Our providers never do.

Stored in the EU

Your files, database and our servers are in Paris, and our main speech-recognition provider works in the EU. The providers that process elsewhere are listed below.

Deleted when you say so

Deleting a project removes the original and every file made from it. Bulk exports delete themselves after a day.

Copyright and ownership

You keep full ownership of everything you upload: audio, video, scripts and other material. Everything the service produces from it is yours too, including transcripts, subtitles, translations, chapters and rendered videos, provided you hold the rights to the original.

We claim no rights to your content. The only permission you give us is what we need to run the service for you: storing your files, processing them, making preview copies and

thumbnails, and backing up our database. We never publish your content or share it with anyone unless you ask us to, or the law requires it.

You are responsible for having the rights to upload what you upload. This is set out in articles 1 and 5 of our Terms of Service.

AI training

We only use your content to train, fine-tune or evaluate AI models, or to tune our prompts and rules, if your organisation has agreed to it in writing. Without that agreement, your content is used only to run the service for you.

We use AI providers to transcribe and translate. Each one processes your content only to return a result to us, and none may use it to train or improve its models:

- **Soniox**, our main speech-recognition provider, contractually never uses customer audio or transcripts to improve its models.
- **Reson8**, which transcribes files uploaded as Dutch – dialect, does not train on customer data.
- **ElevenLabs** and **Speechmatics**, our backup speech-recognition providers, are set up so they do not use our customers' data for training.
- **Google** provides the Gemini language models through its paid API, where Google does not use prompts or responses to improve its products.

What happens when you upload a file

1 **Upload**

Your browser sends the file straight to our storage in Paris over an encrypted connection, using a one-time upload link that expires after an hour.

2 **Preparation**

We make a 720p preview copy for smooth playback, a waveform and a strip of thumbnails. They are stored next to the original.

3 **Transcription**

A speech-recognition provider fetches the file and returns a transcript. That is Soniox, on its EU endpoint, unless it is unavailable, in which case ElevenLabs or Speechmatics takes over. Soniox deletes the job as soon as we have the transcript. Files uploaded as Dutch – dialect go to Reson8 in the Netherlands first, which keeps no audio, and to Soniox if Reson8 is unavailable.

4 **Subtitles and translation**

Google's Gemini models receive the transcript text, never the audio or video, to lay out subtitles, translate them, and suggest chapters and highlights. The results are saved in our database in Paris.

5 **Editing**

You edit in your browser and changes are saved to our database. To load quickly, the editor keeps a copy of your transcript in your browser's storage; on a shared computer, clear the site's data when you are done.

6 **Rendering and export**

Videos with burned-in subtitles are rendered on our servers in the EU. The finished file is private: it downloads through a short-lived link, and only after we have checked that you have access to the project.

Who can access your content

- **You and the people you choose.** Within your organisation, a project is visible to the members you give access to it or to its folder. Organisation admins do not automatically see every project.
- **People with a file link.** Your media files, such as the original upload and its preview copy, are delivered through links that do not require signing in. We only show those

links to people with access to the project, but anyone you forward one to can open that file. We are replacing them with links that expire.

- **Our team.** We do not browse customer content. A small number of engineers can technically reach it, and they open a project only when you ask us for help with it, or when it is needed to keep the service secure or to comply with the law.
- **Our providers.** Each receives only what its task needs. The sub-processor list sets out exactly what and where.
- **Authorities.** We disclose content only when the law compels us to, and we tell you first unless the law forbids it.

Where your content is stored and processed

Files	Amazon Web Services, Paris
Database	MongoDB Atlas, on Amazon Web Services in Paris
Web application	Vercel, server functions in Paris
Processing and rendering	Amazon Web Services, Paris
Speech recognition	Soniox (EU); Reson8 (the Netherlands) for Dutch – dialect; backups Speechmatics (EU) and ElevenLabs (United States, moving to the EU)
Language models	Google Gemini API (location not fixed)

Where content leaves the European Economic Area, the transfer is covered by an adequacy decision or by the European Commission's Standard Contractual Clauses.

How long we keep it

- **Until you delete it.** Projects are never deleted automatically, on any plan.
- **Deleting a project** permanently removes the original file, its preview copy, waveform and thumbnails, its transcripts and translations, and its rendered videos, and clears cached copies from our delivery network. If anything cannot be deleted, you see an error and the project stays, so you can try again; it is never left half-deleted.
- **Bulk exports** are deleted automatically one day after they are created.
- **Deleting your account** removes the organisations you own, with all their projects and files. Billing records are kept as long as accounting law requires.
- **Backups.** Deleted data can remain in encrypted database backups until those backups expire.
- **Providers.** Soniox deletes each transcription job as soon as we have the result, and anything left after 30 days automatically. Reson8 keeps no audio. Speechmatics deletes jobs after 7 days. Google may keep Gemini requests for a limited period to detect abuse.

How we protect it

- Every connection is encrypted with HTTPS, which browsers are told to enforce.
- Files and the database are encrypted at rest with AES-256, by Amazon S3 and MongoDB Atlas.
- Every page and every request is checked against your sign-in and your access to the project.
- Passwords are stored as bcrypt hashes. Email addresses are verified, and sign-in and sign-up are rate-limited and protected against bots.
- Uploads that a browser could execute, such as HTML or SVG files, are refused.

Documents for your compliance team

- Terms of Service
- Privacy Policy
- Data Processing Agreement
- Sub-processors

Every page can be saved as a PDF. If your organisation needs something we have not covered here, such as answers to a security questionnaire, email privacy@subtitle.ai.