

Nou servei de transcripció d'àudio i vídeo en català

written by David Folch | 8 de març de 2023

Després d'unes setmanes en proves, Softcatalà presenta oficialment el **servei de transcripció d'àudio i vídeo a text**. Aquest servei, també anomenat dictat, es troba [disponible al web de Softcatalà](#).

Amb aquest nou servei serà més senzill transcriure els fitxers en català d'àudio i vídeo a text, sense necessitat d'eines en altres llengües. Es basa en models lliures de **xarxes neuronals** entrenades per a processar els fitxers i generar el text corresponent.

Transcripció de fitxers a text

Podeu enviar-nos un fitxer d'àudio o vídeo. Us avisarem per correu quan estigui transcrit i el podreu passar a recollir. Hi ha una cua de fitxers que anem processant. Pot trigar més o menys segons el nombre de transcripcions pendents de processar.

Model que voleu utilitzar:

Correu on voleu que us avisem:

Selecioneu un fitxer d'àudio o vídeo

No s'ha triat cap fitxer

Mida màxima del fitxer: 1GB. Formats admesos: mp3, wav, ogg, flac, avi, mp4, mov i mkv

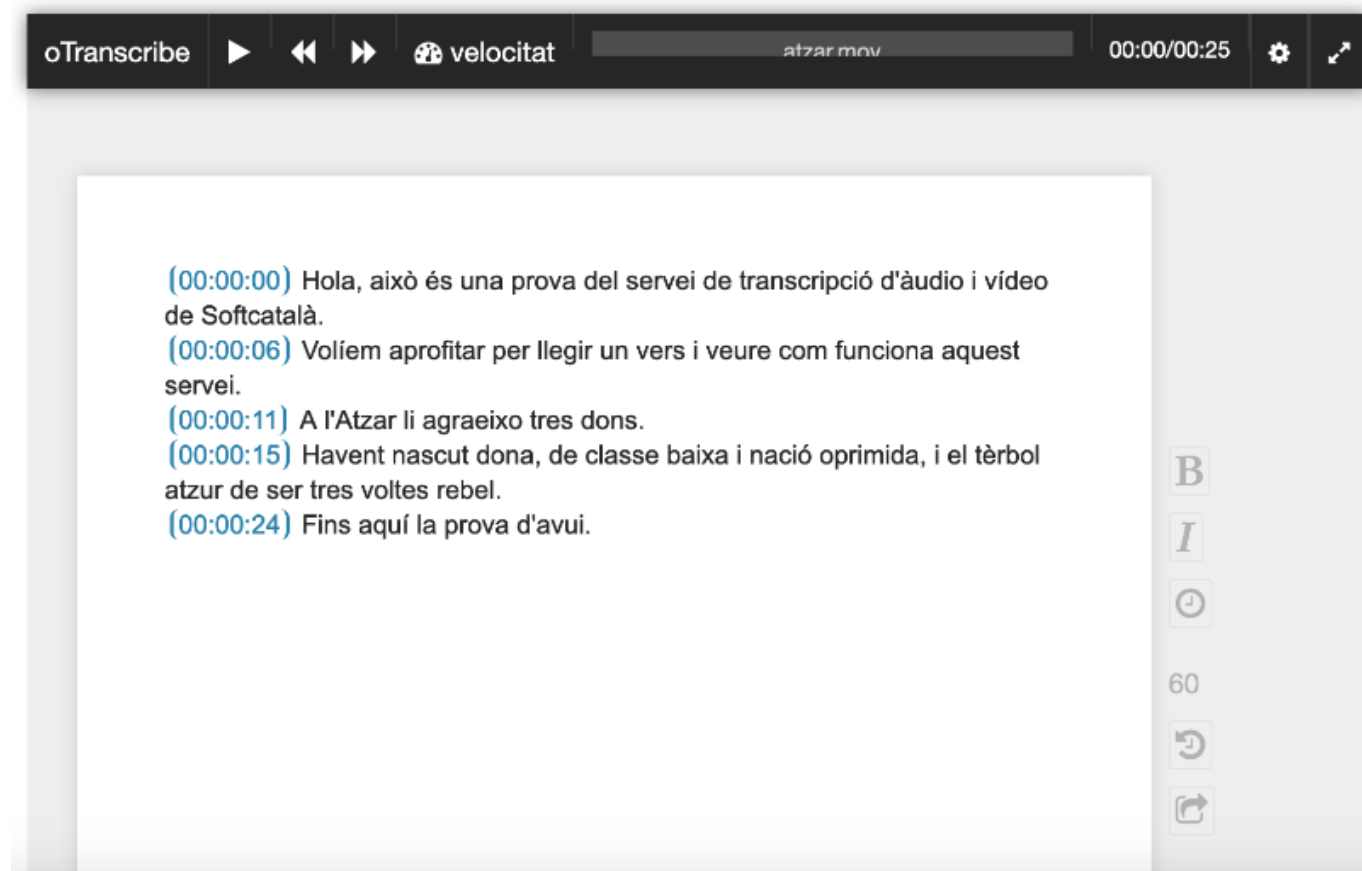
Com s'utilitza el servei de transcripció?

El servei de transcripció és molt fàcil d'utilitzar. Només cal enviar un fitxer d'àudio o vídeo en algun dels següents formats: MP3, WAV, OGG, FLAC, AVI, MP4, MOV o MKV.

Aquest fitxer es processarà als nostres servidors i, quan finalitze aquesta acció, s'enviarà un correu electrònic a l'usuari amb l'enllaç a la transcripció. La duració d'aquest procés, o el temps d'espera, depèn del nombre de transcripcions que hi haja pendents de processar. **El fitxer s'esborrarà dels nostres sistemes 72 hores després de ser transcrit.**

Un cop finalitzada la transcripció, es pot baixar el text de la transcripció tant en format text com en SubRip (str), que és el format més popular de subtítols.

A més de baixar els textos, també hi ha l'opció d'un editor en línia que permet acabar d'esmenar la transcripció, ja que **aquests serveis cometten errades**. Com ocorre amb qualsevol servei automàtic, sempre cal que els usuaris revisen els resultats per a garantir que no s'hi han comès errades.



The screenshot shows the oTranscribe web interface. At the top, there is a dark control bar with the text 'oTranscribe' on the left, followed by play, stop, and next buttons, a volume icon, the word 'velocitat' (speed), a progress bar labeled 'atzar.mov', and a timer showing '00:00/00:25'. On the right of the bar are settings and share icons. The main content area is a white box containing a transcript with blue timestamps and black text. The transcript reads: '(00:00:00) Hola, això és una prova del servei de transcripció d'àudio i vídeo de Softcatalà. (00:00:06) Volíem aprofitar per llegir un vers i veure com funciona aquest servei. (00:00:11) A l'Atzar li agraeixo tres dons. (00:00:15) Havent nascut dona, de classe baixa i nació oprimida, i el tèrbol atzur de ser tres voltes rebel. (00:00:24) Fins aquí la prova d'avui.' On the right side of the transcript box, there is a vertical toolbar with icons for bold (B), italic (I), a clock icon, the number '60', a refresh icon, and a share icon.

Informació addicional

El servei que presenten transcriu el text del fitxer d'àudio/vídeo, però també la puntuació i les majúscules, cosa que fins ara cap altre sistema oferia.

Per a poder oferir aquest servei, utilitzen els models de transcripció lliures de l'empresa OpenAI, anomenats [Whisper](#). Softcatalà té [un projecte](#) focalitzat en millorar-los per al català.

QUANT A SOFTCATALÀ

Softcatalà és una associació sense afany de lucre que treballa per la normalització de la llengua catalana en el sector informàtic, Internet i les noves tecnologies. Per a més informació sobre els projectes de l'associació, podeu consultar el seu lloc web a <https://www.softcatala.org/>.

(extret del web)