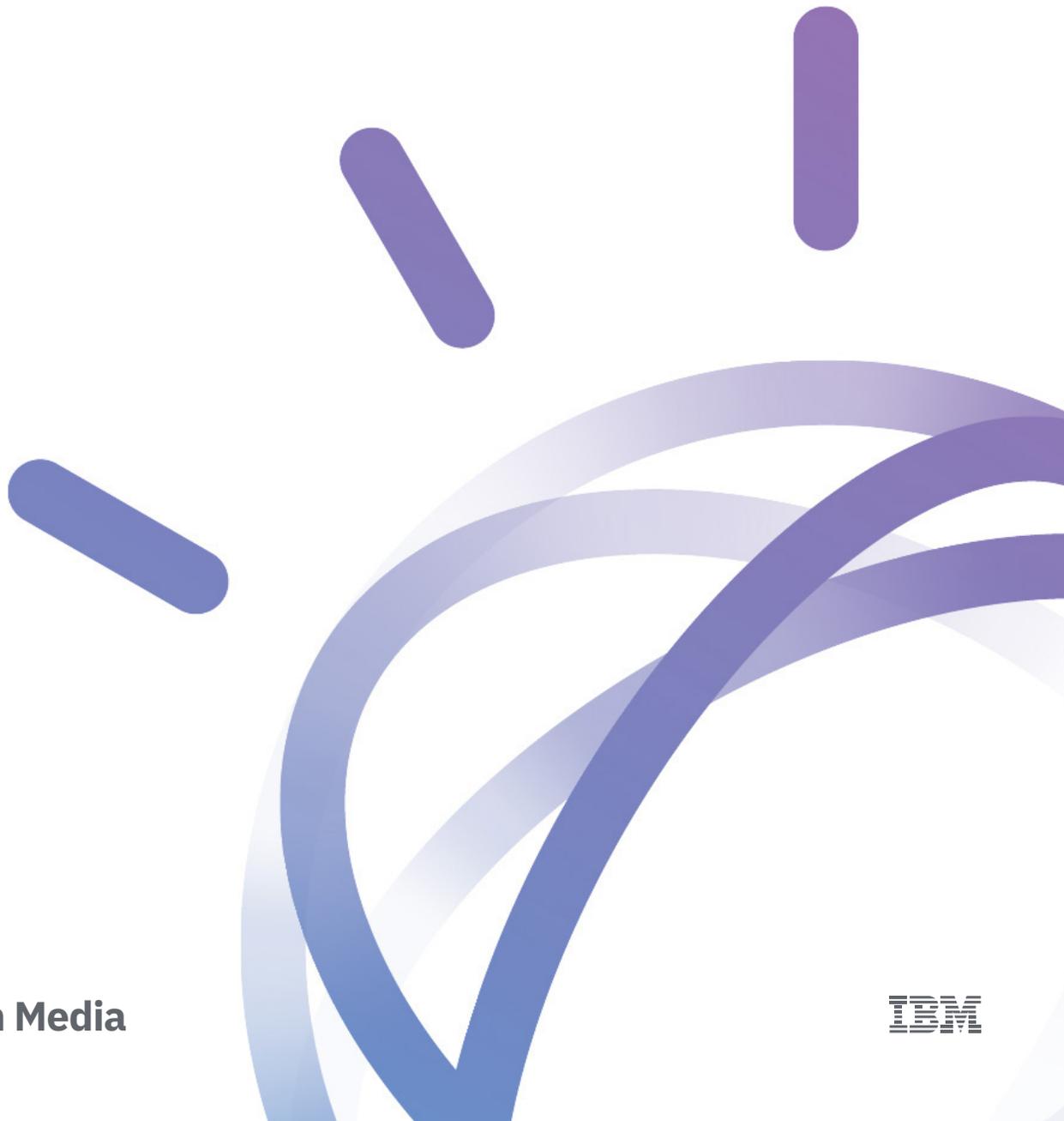# Watson Captioning

IBM

IBM

Watson Captioning (WC), the second offering in a suite of IBM Watson Media solutions, uses Watson's cognitive capabilities to automate transcription and closed captioning services – providing businesses a scalable solution that saves money and time. Watson Captioning streamlines workflows to maximize productivity and uses machine learning technology to increase caption accuracy over time. Rapid generation of accurate, editable captions also makes it easier to meet the changing legal and compliance standards of entities such as the FCC and the ADA.

The Process
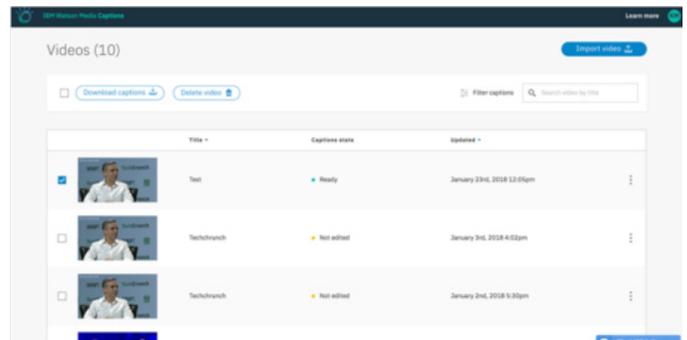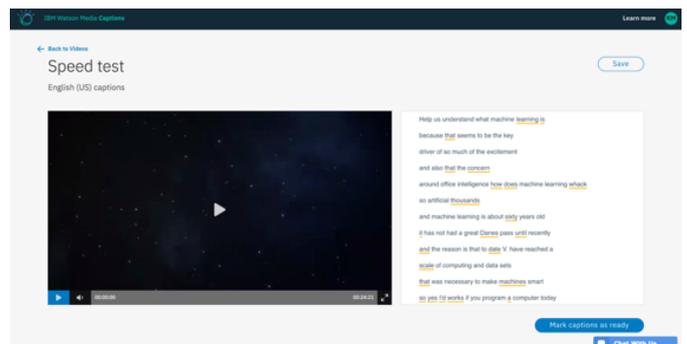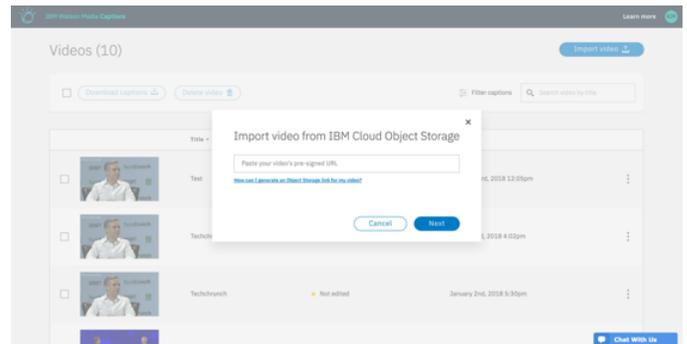
IMPORT & CAPTION GENERATION

Watson Captioning begins with content ingest. After an asset is ingested into the product, via file upload or URL, caption generation begins. Watson watches the video and using the Speech to Text API, transcribes the entire asset at a pace faster than real time. Using a smart layout algorithm, Watson Captioning automatically segments caption cues at natural breaking points, resulting in a file with greater readability and in turn, easier editing. After captioning is complete, a plain text file is generated with low confidence words underlined and the asset status in the UI is updated to "Not Edited". Because Watson Captioning is self-learning, accuracy rates will only improve over time and low confidence words will be seen less often.

CAPTION EDITOR & QUALITY CONTROL

The video player matches the cursor location in the plain text file - meaning that the asset will automatically resume play from the cue line that the user has selected. Using the product's caption editor interface, clients are able to manually correct errors and insert line breaks directly in the caption file. Based on the timestamps of the modifications made, the necessary time synchronization adjustments are automatically taken care of. Users are able to make all edits at one time or save and exit the product with captions partially complete. When the client is satisfied with the final caption file, the asset is marked as "Ready".
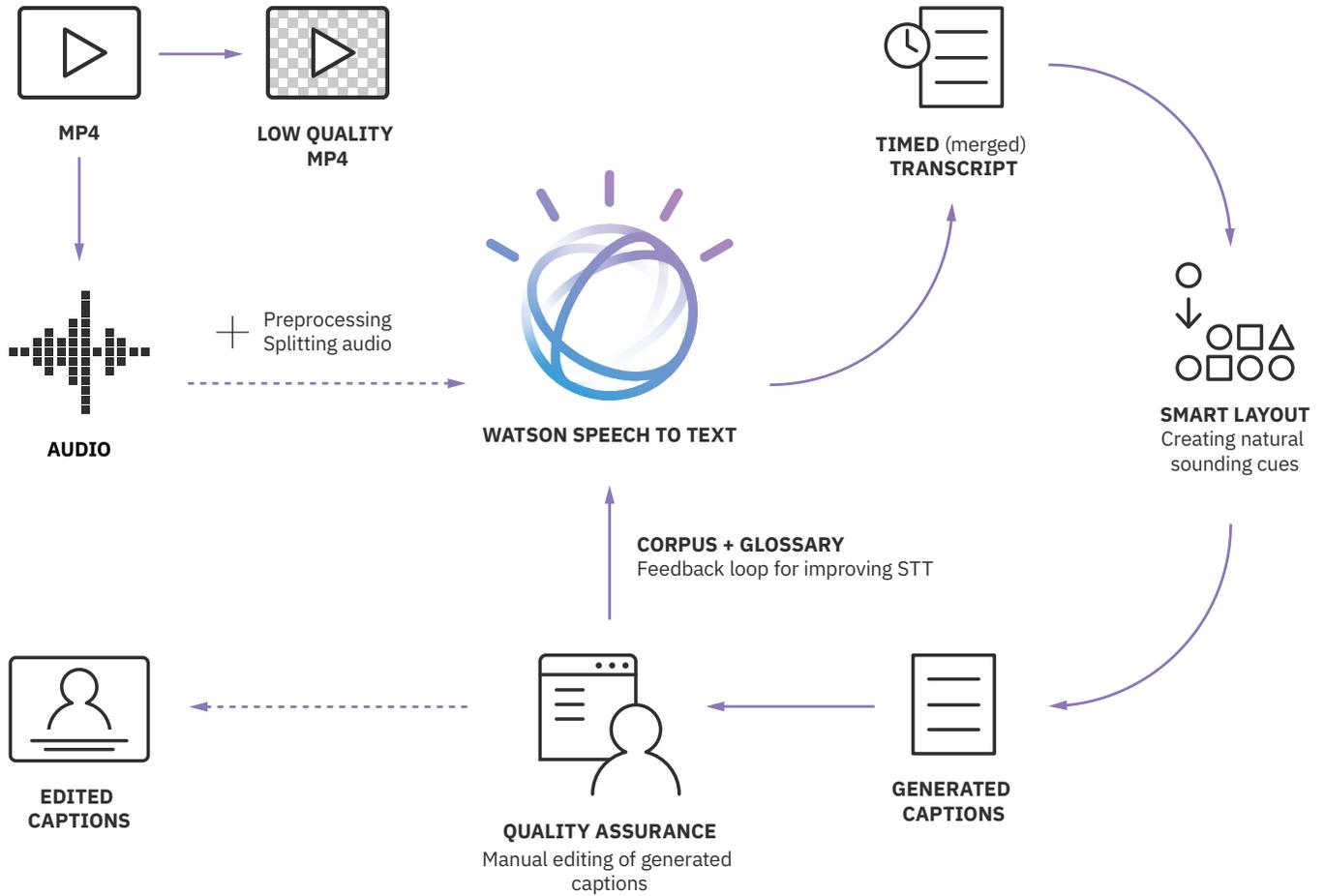

DOWNLOAD & EXPORT

Users are able to download one file directly by clicking the three dots on the right-hand side of an asset line. For a mass export, select the files using the left column check box and click the "Download Captions" button at the top of the home page. All exports are delivered in the form of a zip file. Exports with multiple captions files are broken down by asset within one zip.

IBM Watson Media
Watson Captioning

How It Works

MP4

LOW QUALITY MP4

AUDIO

+ Preprocessing
Splitting audio

WATSON SPEECH TO TEXT

TIMED (merged) **TRANSCRIPT**

SMART LAYOUT
Creating natural sounding cues

CORPUS + GLOSSARY
Feedback loop for improving STT

GENERATED CAPTIONS

QUALITY ASSURANCE
Manual editing of generated captions

EDITED CAPTIONS

To learn more about Watson Captioning, visit www.ibm.com/watson/media/watson-captioning.

IBM Watson Media