GLOBAL ALLIANCE OF SPEECH–TO–TEXT CAPTIONING GUIDELINES FOR REQUESTING CAPTIONING SERVICES

PURPOSE OF DOCUMENT:

This document will aid caption consumers and those requesting captioning services to:

- understand the various forms and methods of captioning
- choose the appropriate method of captioning for an event
- request appropriate caption display options
- consider technical support for captioner/service provider
- consider prep materials for captioner/service provider

DEFINITIONS:

ONSITE CAPTIONING: The captioner/service provider is in the same location as your event.

REMOTE CAPTIONING: The captioner/service provider is listening from a remote location and streaming captions to an online platform or encoder.

REALTIME CAPTIONING: Live, simultaneous display of text to match the spoken word, and may include parenthetical information to alert the consumer to nonverbal sounds or actions.

CART CAPTIONING: Communication access realtime translation (CART) is the same as realtime captioning. At one time, CART captioning was differentiated from broadcast captioning, which may use an encoder to display text.

TELEPHONE CAPTIONING: Captioning that allows you to talk on the phone while reading captions of the conversation. Captions are displayed on a screen on the telephone.

STENO CAPTIONING: A term used to describe an individual trained to produce realtime text from the spoken word using a specialized machine called a steno machine. The steno captioner uses combinations of keys on the machine that correspond to the phonetics of a word or phrase, which is called machine shorthand. The machine shorthand is translated instantaneously to readable text by a computer-aided transcription software that is programmed by the steno captioner and that can then be output to various display modalities.
"Steno" is a shortened form of the word "stenography," which is the process of writing in shorthand.

VOICE WRITING: (AKA subtitles or respeaking) A term used to describe an individual trained to use their voice and respeak into a microphone and/or stenomask using software utilizing automatic speech recognition (also referred to as ASR). The voice writer provides speech-to-text in realtime by listening and respeaking what an individual(s) is saying, which instantaneously translates via a computer-aided program into text and can be streamed using different software programs.

ASR CAPTIONING: Automatic speech recognition (ASR) captioning is the process of converting speech to text by means of algorithms implemented by a computer device and then visually displaying that text.
ASR captioning can be realtime or postproduction and can utilize various methods to recognize and translate the spoken word into text. ASR technology has benefited recently by advancements in artificial intelligence (AI) and machine learning (ML) techniques.

TEXT INTERPRETING: The text interpreting captionist inputs text using a keyboard abbreviation system based on phonetics and produces text of spoken information in various venues providing summary and/or meaning-for-meaning translation of the spoken word. Examples are C-Print and Typewell.

POSTPRODUCTION CAPTIONING: Captioning that is prepared from recorded media once that media has been edited and finalized.

ON-DEMAND VIDEO CAPTIONING: Captioning that appears when a video on demand (VOD) on the Internet is accessed. Captions for on-demand video may be created using any of the following methods: postproduction captioning/subtitling, realtime or ASR captioning.

CONSIDERATIONS WHEN REQUESTING CAPTIONING

Captions should be accurate, consistent, clear, readable and accommodate the needs of the audience. Live captioning varies depending on the needs and expectations of the audience. Captioning is often provided to comply with accessibility regulations affecting persons with hearing loss. Captioning may also showcase automatic speech recognition technology, create searchable text of an event, or improve comprehension, particularly by speakers of other languages.

VENUE:
- Where will your event be held?
- What is the format of the event?
- Meeting with multiple speakers

- Single speakers from podium
- Q/A interview style
- Roundtable/panel discussions
- Audience participation
- Breakout sessions
- How will people attend?
- How many people are expected to attend?
- Will the event be streamed or broadcast live?
- Will slides be used?
- Will videos be shown? Are the videos captioned?

TECHNICAL:

- How will the captioner/service provider get audio?
- Is there a sound system that will be used?
- Will there be Wi-Fi available in the venue?
- Will professional audio/visual personnel be present?
- Is event being recorded? Will recordings be captioned?
- How will audience members view the captioning?
- Will video of the event be displayed in the venue?
- Should captions appear on the same video screen?

PREPARATORY MATERIALS:

- Will your event contain technical material/terminology?
- Will scripts be used?
- Who will get prep to the captioner/service provider?
- Is there a website for the event?