

**Before the
Federal Communications Commission
Washington, D.C. 20554**

In the Matter of:)	
Closed Captioning of Video Programming)	CG Docket No. 05-231
Telecommunications for the Deaf and)	MB Docket No. RM-11065
Hard of Hearing, Inc. Petition for)	Docket No. RM-_____
Rulemaking)	

Petition for Declaratory Ruling and/or Rulemaking on Live Closed Captioning Quality Metrics and the Use of Automatic Speech Recognition Technologies

**Telecommunications for the Deaf and Hard of Hearing, Inc. (TDI)
National Association of the Deaf (NAD)
Hearing Loss Association of America (HLAA)
Association of Late-Deafened Adults (ALDA)
Cerebral Palsy and Deaf Organization (CPADO)
Deaf Seniors of America (DSA)
Deaf/Hard of Hearing Technology Rehabilitation
Engineering Research Center (DHH-RERC)
Twenty-First Century Captioning Disability and Rehabilitation Research Project
(Captioning DRRP)
Rehabilitation Engineering Research Center on Universal Interface &
Information Technology Access (IT-RERC)
National Technical Institute for the Deaf**

via electronic filing
July 31, 2019

Samuelson-Glushko Technology Law & Policy
Clinic (TLPC) at Colorado Law

Counsel to TDI

Blake E. Reid, Director

blake.reid@colorado.edu

Telecommunications for the Deaf and Hard of Hearing, Inc. (TDI)

Claude Stout, Executive Director • cstout@TDIforAccess.org
PO Box 8009, Silver Spring, MD 20907
www.TDIforAccess.org

National Association of the Deaf (NAD)

Howard Rosenblum, Chief Executive Officer • howard.rosenblum@nad.org
Contact: Zainab Alkebsi • zainab.alkebsi@nad.org
8630 Fenton Street, Suite 820, Silver Spring, MD 20910
301.587.1788
www.nad.org

Hearing Loss Association of America (HLAA)

Barbara Kelley, Executive Director • bkelly@hearingloss.org
Lise Hamlin, Director of Public Policy, LHamlin@Hearingloss.org
7910 Woodmont Avenue, Suite 1200, Bethesda, MD 20814
301.657.2248
www.hearingloss.org

Association of Late-Deafened Adults (ALDA)

Richard Brown, President • President@alda.org
8038 MacIntosh Lane, Suite 2, Rockford, IL 61107
815.332.1515
www.alda.org

Cerebral Palsy and Deaf Organization (CPADO)

Mark Hill, President • president@cpado.org
14510 Homecrest Road Unit # 3008, Silver Spring, MD 20906
503.512.5066
www.cpado.org

Deaf Seniors of America (DSA)

Nancy B. Rarus, President • nbrarus@gmail.com
Alfred Sonnenstrahl, Vice President • alsonny@icloud.com
5619 Ainsley Court, Boynton Beach, FL 33437

Rehabilitation Engineering Research Center on Technology for the Deaf and Hard of Hearing, Gallaudet University (DHH-RERC)

Twenty-First Century Captioning Disability and Rehabilitation Research Project (Captioning DRRP)

Christian Vogler, PhD • christian.vogler@gallaudet.edu
800 Florida Avenue NE, TAP – SLCC 1116, Washington, DC 20002

**Rehabilitation Engineering Research Center on Universal Interface & Information
Technology Access (IT-RERC)**

Gregg Vanderheiden, PhD, Director • greggv@umd.edu

Trace Research & Development Center • University of Maryland
4130 Campus Drive, College Park, MD 20742

National Technical Institute for the Deaf

Dr. Gerard Buckley, President • gjbcfo@ntid.rit.edu

Gary Behm, VP of Academic Affairs • gwbnts@rit.edu

52 Lomb Memorial Drive, Rochester, NY 14623

www.ntid.rit.edu

Summary

For more than two decades, the quality of closed captions of live programming has stood as a significant but underdeveloped priority for the accessibility of video programming for Americans who are deaf or hard of hearing. While the Commission has acknowledged the critical importance of accurate, synchronous, complete, and well-placed captions to ensure equal access to news, weather, sports, and other live programming, its focus on “best practices” for captioning methodology has failed to yield consistent results. As new captioning technologies and methodologies including automatic speech recognition (ASR) enter the captioning marketplace, many consumers have continued to experience poor-quality captions on live programming, which in some cases have become even worse over the past several years.

It is time for a change. In this petition, we urge the Commission to finally begin in earnest an inquiry aimed at developing objective, technology-neutral metrics for caption quality. While developing these metrics remains a difficult task, a “best practices” approach tailored to traditional methods of human and ENT captioning is not a workable approach for the diverse, modern landscape of live captioning methodologies and technologies that increasingly incorporate automation and other approaches with widely varying results.

More specifically, we urge the Commission to initiate an inquiry into the state of the art of closed captioning techniques for live television programming and how the varying dimensions of caption quality, including accuracy, synchronicity, completeness, and placement affect the accessibility of video programming. Following the development of a robust record, we urge the Commission to turn to a rulemaking to require live television programming to be captioned at a level that meets or exceeds technology-neutral metrics calibrated to guarantee that the programming is accessible by Americans who are deaf or hard of hearing. Finally, we urge the Commission to address near-term issues with the use of ASR by immediately issuing a declaratory ruling and/or expedited rule change with near-term guidance and policy on the application of the existing best practices to ASR.

Table of Contents

Summary	iv
Discussion.....	1
I. The Commission did not grapple seriously with quality problems in the provision of captions for live programming for nearly two decades.	2
II. The human- and ENT-specific “best practices” approach in the <i>2014 Caption Quality Order</i> does not assess or ensure the <i>quality</i> of captions.....	5
III. Quality problems with captions for live programming persist widely under the “best practices” regime.	10
IV. The human and ENT-centric, quality “best practices” approach provides no means for objectively assessing how increasingly diverse captioning methodologies compare and whether they result in accessible video programming.	13
V. The Commission should launch an inquiry to develop objective metrics for caption quality and processes for enforcing them.	14
VI. The Commission must provide immediate guidance to video programmers on the permissible use of automatic speech recognition.....	16

Discussion

Pursuant to Rules 1.2, 1.401, and 1.412(c),¹ the above-listed organizations (“Consumer Groups”) and accessibility researchers respectfully petition the Commission to:

- a. Initiate an inquiry into the state of the art of closed captioning techniques for live television programming and how the varying dimensions of caption quality, including accuracy, synchronicity, completeness, and placement affect the accessibility of video programming;
- b. Use the record to develop rules requiring live television programming to be captioned at a level that meets or exceeds technology-neutral metrics guaranteeing that the programming is accessible by Americans who are deaf or hard of hearing;
- c. Immediately issue a declaratory ruling and/or expedited rule change with near-term guidance and policy on the use of automatic speech recognition (ASR) technologies for captioning of live television programs.²

Consumer Groups seek to promote equal access to video programming for the more than 48 million Americans who are deaf, hard of hearing, late-deafened, or DeafBlind so they may fully experience the informational, educational, cultural, and societal opportunities afforded by the telecommunications revolution. Toward these ends, the accessibility researchers conduct basic and applied research into the accessibility of technology to consumers with disabilities, including a five-year project by the Captioning DRRP solely dedicated to video programming.

The closed captioning provisions of the Telecommunications Act of 1996 require the Commission to “ensure that . . . video programming . . . is *fully* accessible through the provision of

¹ 47 C.F.R. §§ 1.2, 1.401, 1.412(c). Given the large number of potential commenters, including individual consumers, on the issues raised in this petition, pursuant to Rule 1.47(d) the petitioners consent to electronic service via ECFS of comments in support of or opposition to this petition to the extent that service is required under Rule 1.405(a) or any other relevant rule, and urge the Commission to allow electronic filing of comments and replies on this petition to the greatest extent possible when placing this petition on public notice.

² 47 C.F.R. § 79.1(j)(2)-(3). This request is described in greater detail *infra*, Part VI.

closed captions.”³ In the nearly quarter century since, members of the Consumer Groups and other advocates have repeatedly counseled the Commission that objective, technology-neutral quality standards for live captioning are critical “to prevent a proliferation of low quality captioning services that might otherwise result.”⁴ In that time, the Commission has gestured toward but never implemented such metrics, leaving quality largely unaddressed for the first two decades of the closed captioning rules and in the last five years imposing a methodology-specific “best practices” approach that has not resulted in sufficient improvements in quality to serve the Commission’s legal mandate.

Accordingly, we urge the Commission to immediately begin an inquiry into the development of technology-neutral quality metrics that evaluates captions on the extent to which they succeed in making video programming accessible. In the meantime, we urge the Commission to address the unchecked proliferation of untested ASR technologies by issuing guidance on the applicability of existing best practices to the use of ASR.

I. The Commission did not grapple seriously with quality problems in the provision of captions for live programming for nearly two decades.

The Commission has long recognized in principle the importance of high-quality closed captioning for live programming to vindicating the civil rights of Americans who are deaf or hard of hearing.⁵ In proposing its initial set of closed captioning rules in 1997, the Commission

³ P.L. 104-104 § 305 (codified at Section 713(a)-(e), (g)-(h) (47 U.S.C. § 613(a)-(e), (g)-(h))).

⁴ *E.g.*, Comments of NAD at 23-24, MM Docket No. 95-176 (Feb. 27, 1997), <https://www.fcc.gov/ecfs/filing/178587>; *see also* Comments of Consumer Action Network (Feb. 27, 1997) (noting that “[c]aptions are of limited use if they are replete with mistakes in spelling, grammar, timing, or placement” or “do not include all of the elements of the soundtrack necessary for accessibility” and observing the “likel[i]hood] that captioning companies that do not aspire to produce high quality work will proliferate” in the absence of standards), <https://www.fcc.gov/ecfs/filing/178490>; Comments of ALDA at 8 (Feb. 25, 1997) (noting that absent quality standards, “providers will have little incentive to contract with high quality captioning providers”), <https://www.fcc.gov/ecfs/filing/178344>.

⁵ *See Closed Captioning & Video Description of Video Programming*, Notice of Proposed Rulemaking, MM Docket No. 95-176, 12 FCC Rcd. 1044, 1090, ¶ 111 (1997) (proposing rules to implement 47 U.S.C. § 613(b)).

acknowledged “that the quality of captioning is a matter of considerable importance to those viewing captions,” “recogniz[ing] that captions must provide information substantially equivalent to that of that of the audio portion of a video program in order to be useful and ensure accessibility.”⁶ As early as 1995, the Commission acknowledged that quality concerns were particularly acute “for live programming where there is no chance to review and correct for errors.”⁷

Despite acknowledging the critical importance of live caption quality, the Commission has kicked the can down the road for nearly two decades on implementing objective quality requirements. That pattern began with its decision not to adopt non-technical caption quality standards in the initial *1997 Captioning Order*.⁸ Likewise, the Commission concluded in the *1997 Order* that it would allow for the use of incomplete, low-quality electronic newsroom (ENT) captioning for live local news programming.⁹

The Commission purported to demur on quality standards only to “allo[w] video programming providers to establish quality standards . . . through their arrangements with captioning suppliers” and promised that it would “revisi[t] this issue if . . . it bec[ame] apparent that [its] assumptions regarding the marketplace incentives for quality [were] incorrect.”¹⁰ It similarly acknowledged concerns that “portions of live newscasts often remain uncaptioned . . . with the use of [ENT]” and committed to later “revisit[ing] this issue . . . to evaluate whether [ENT] provides sufficient captioning of news programming.”¹¹

The concerns over ENT were so significant that the Commission quickly changed course mere months after the *1997 Order*, concluding on reconsideration that ENT was incapable of fully

⁶ *Id.* The Commission acknowledged these concerns as early as 1995, where it noted

⁷ *Closed Captioning & Video Description of Video Programming*, Notice of Inquiry, MM Docket No. 95-176, 11 FCC Rcd. 4912, 4927-28, ¶33 (1995).

⁸ *See Closed Captioning & Video Description of Video Programming*, Report and Order, MM Docket No. 95-176, 13 FCC Rcd. 3272, 3374, ¶ 222 (1997) (“*1997 Captioning Order*”).

⁹ *Id.* at 3311-12, ¶ 84.

¹⁰ *Id.* at 3374, ¶ 222.

¹¹ *Id.* at 3311-12, ¶ 84.

captioning significant portions of news broadcasts and prohibiting the use of ENT by broadcast stations in the top 25 markets and many non-broadcast networks accordingly.¹² In 2000, the Commission acknowledged that the quality problems with ENT were so significant that they “threatened [the] safety” of people who are deaf or hard of hearing and ordered even those entities still allowed to use ENT under normal circumstances to caption or otherwise visually present emergency information.¹³

However, in 2004, nearly a decade after the Commission’s initial adoption of captioning rules, quality problems with captions for live programming, including those generated with ENT, remained so significant that several of the Consumer Groups petitioned the Commission to act.¹⁴ The *2004 Petition* noted that “the quality of captioning generally ha[d] not improved,” and that “[t]same types of captioning quality problems . . . that occurred in 1995 when the Commission opened its [captioning inquiry] continue[d] to occur.”¹⁵ In response to complaints, broadcasters had even begun to explicitly disclaim responsibility for the quality of captions on-air.¹⁶ The *2004 Petition* also noted that quality problems continued to persist with ENT-captioned live programming.¹⁷

The Commission acknowledged the issues surrounding live captioning the raised in the *Petition* by releasing a Notice of Proposed Rulemaking in 2005 that considered the possibility of non-

¹² *Closed Captioning & Video Description of Video Programming*, Order on Reconsideration, MM Docket No. 95-176, 13 FCC Rcd. 19973 ¶¶ 35, 38 (1998) (“*1998 Recon Order*”); see also *Closed Captioning and Video Description of Video Programming*, Clarification Order, MM Docket No. 95-176, 16 FCC Rcd. 5067 (2001) (clarifying the interoperation of the closed captioning rules with ENR captioning during the rules’ initial transition period).

¹³ *Closed Captioning & Video Description of Video Programming, Accessibility of Emergency Information*, Second Report and Order, MM Docket No. 95-176, 15 FCC Rcd. 6615, 6620-21, 6623-24, ¶¶ 10, 16 (2000)

¹⁴ Petition for Rulemaking of TDI, et al., MB Docket No. RM-11065 (Aug. 20, 2004), <https://www.fcc.gov/ecfs/filing/5511440137>.

¹⁵ *Id.* at 35

¹⁶ *Id.* at 27

¹⁷ *Id.* at 34-35.

technical quality standards, including for live programming, and extending the prohibition of ENT outside the top 25 markets.¹⁸

Despite receiving more than 1600 filings in response to its request for comments on the *2004 Petition* the Commission took no action on the *Petition's* concerns about live captioning, and the Consumer and Governmental Affairs Bureau concluded five years later in 2010 that the record on quality had gone stale.¹⁹ The Bureau again sought comment on the quality issues surrounding live programming, including the continued use of ENT.²⁰

II. The human- and ENT-specific “best practices” approach in the *2014 Caption Quality Order* does not assess or ensure the *quality* of captions.

Almost two decades after originally acknowledging the potential for quality issues with the captioning of live programming, the Commission finally took action in 2014 on non-technical caption quality issues, including for live programming.²¹ In the *2014 Caption Quality Order* the Commission concluded that its “original assumptions regarding the marketplace incentives for quality captioning ha[d] not been borne out.”²² The Commission cited “[t]he lack of regularity in the quality of closed captions” and “the continuing dissatisfaction reported in hundreds of individual accounts submitted to the Commission” as evidence of “widespread frustration among the viewing public with inconsistencies in caption quality.”²³ The Commission noted that “a substantial number

¹⁸ *Closed Captioning of Video Programming, Telecommunications for the Deaf & Hard of Hearing, Inc. Petition for Rulemaking*, Notice of Proposed Rulemaking, 20 FCC Rcd. 13,211, 13,217, 13,226-27 ¶¶ 13-16, 48 (2005).

¹⁹ *See Consumer & Governmental Affairs Bureau Seeks to Refresh the Record on Notices of Proposed Rulemaking Regarding Closed Captioning Rules*, Public Notice, CG Docket No. 05-231, ET Docket No. 99-254, 25 FCC Rcd. 15,056, 15,056-57 (2010).

²⁰ *Id.* at 15,057-58.

²¹ *Closed Captioning of Video Programming, Telecommunications for the Deaf & Hard of Hearing, Inc. Petition for Rulemaking*, Report and Order, Declaratory Ruling, and Further Notice of Proposed Rulemaking, 29 FCC Rcd. 2221, at 2223-24, ¶¶ 2-3 (2014) (“*2014 Caption Quality Order*”).

²² *Id.* at 2237-38, ¶ 22.

²³ *Closed Captioning of Video Programming, Telecommunications for the Deaf & Hard of Hearing, Inc. Petition for Rulemaking*, Report and Order, Declaratory Ruling, and Further Notice of Proposed Rulemaking, 29 FCC Rcd. 2221, 2237-38, ¶ 22 (2014) (“*2014 Caption Quality Order*”).

of consumer commenters [had] described various problems with caption quality . . . , including captions that are inaccurate, gibberish, garbled, butchered, incomplete, misspelled and/or misunderstood, incomprehensible, obscuring the speaker, or significantly lag behind the spoken words they are intended to convey” and that “the refreshed record demonstrate[d] that the consumer experience has not improved during the intervening years.”²⁴ The Commission also acknowledged the Consumer Groups’ ongoing concerns that the use of ENT to generate poor-quality captions had continued to deny viewers access to news programming in markets across the country.²⁵

As a result, the Commission adopted overarching quality standards. The standards require captions to “convey the aural content of video programming . . . to individuals who are deaf or hard of hearing to the same extent that the audio track conveys such content to individuals who are able to hear,” and be “accurate, synchronous, complete, and appropriately placed.”²⁶

However, the Commission adopted a weakened, multi-factor, case-by-case *de minimis* standard for evaluating captioning errors for live and near-live programming.²⁷ This weakened standard has left open the door for significant quality errors in the captioning of live programming.

More problematically, the Commission declined to adopt Consumer Groups’ proposals to adopt objective, technology-neutral quality metrics for assessing compliance with the caption quality standards.²⁸ These metrics would have held video programmers responsible for the ultimate quality of captions for live programming without regard to the methodology used to create them, encouraging the development of effective techniques to improve the quality and cost of captioning.²⁹

²⁴ *Id.* at 2237, ¶ 21 (internal citations omitted).

²⁵ *Id.* at 2268, ¶ 75.

²⁶ 47 C.F.R. § 79.1(j)(2).

²⁷ 47 C.F.R. § 79.1(j)(3).

²⁸ *Id.* at 2257, ¶ 59.

²⁹ *See id.*

Instead, the Commission chose instead to adopt a set of methodology-specific “best practices” for video programmers,³⁰ captioning vendors,³¹ and real-time (live) captioners³² that do not require any specific, objective level of quality. The best practices instead purport to promote quality for live programming indirectly by allowing video programmers to opt into a check-the-box approach that defers to caption vendors, captioners, and video programmers themselves to make and follow their own rules for caption quality.

More specifically, the best practices require video programmers who opt to satisfy the caption quality standards through best practices³³ to adopt “performance requirements” that are “comparable to” the best practices in their agreements with caption vendors, undergo efforts to verify compliance and ensure training, and make available advance preparation materials and high-quality audio.³⁴ In turn, caption vendors contracting with programmers adopting best practices must “create and use metrics to assess” the quality of real-time (live) captions, “establish minimum acceptable standards based on those metrics,” “striv[e] to regularly exceed those minimum standards,” and “perform frequent and regular evaluations and sample audits to ensure those standards are maintained.”³⁵ Real-time captioners working under best practices, in turn, must caption as “accurately, synchronously, completely, and appropriately placed as possible, given the nature of the programming,”³⁶

³⁰ 47 C.F.R. § 79.1(m)(1)(ii) (allowing video programmers to comply either with the quality standards in 47 C.F.R. § 79.1(j)(2) or the best practices in 47 C.F.R. § 79.1(k)(1)).

³¹ 47 C.F.R. § 79.1(k)(2) (real-time (live) captioning vendors) & (k)(4) (offline (prerecorded) captioning vendors).

³² 47 C.F.R. § 79.1(k)(3).

³³ Under Rule 79.1(m)(1)(ii), video programmers must annually certify either that their programming satisfies the caption quality standards in Rule 79.1(j)(2) or that they follow the best practices in Rule 79.1(k)(1). 47 C.F.R. § 79.1(m)(1)(ii).

³⁴ 47 C.F.R. § 79.1(k)(1)(i)-(ii).

³⁵ 47 C.F.R. § 79.1(k)(2)(i)-(iii).

³⁶ 47 C.F.R. § 79.1(k)(3)(i).

The Commission acknowledged but dismissed the Groups' concerns that "the ultimate quality of captions delivered to consumers, and not the process by which they are created, is the only logical and acceptable metric for the Commission to review" and that "follow[ing] some particular process to create captions for a program cannot cure the program's inaccessibility if the process ultimately results in poor-quality captions."³⁷ Instead, the Commission speculated that best practices would be effective, and committed to reassessing them one year after the effective date of the imposition of the quality standards and revisiting its rules to address any shortcomings.³⁸

The Commission likewise demurred on the Consumer Groups' long-standing proposal to ban ENT across the board.³⁹ The Commission acknowledged the ongoing shortcomings of ENT and Consumer Groups' contention that ENT, which, roughly speaking, converts the content of newscasters' teleprompters to captions, is *per se* methodologically incapable of generating reasonably complete and accurate captions of unscripted programming,⁴⁰

However, the Commission again concluded that the long-running quality issues could instead be addressed by adopting another set of check-the-box best practices for "enhancements" to ENT.⁴¹ These practices require broadcast stations using ENT to increase scripting of in-studio news content,⁴² weather interstitials,⁴³ and pre-produced programming,⁴⁴ include supplemental crawls for live interviews and breaking news,⁴⁵ improve training,⁴⁶ and appoint an "ENT Coordinator."⁴⁷ As with the best practices for real-time programming, the Commission committed to revisiting phasing

³⁷ See *2014 Caption Quality Order*, 29 FCC Rcd. at 2257, ¶ 59 (quoting *Ex Parte of TDI, et al.*, CG Docket No. 05-231 at 6 (Jan. 15, 2014)).

³⁸ See *id.*

³⁹ See *id.* at 2268-69, ¶ 76

⁴⁰ See *id.* at 2267-68, ¶¶ 75-78

⁴¹ See *id.* at 2268-69, ¶ 76.

⁴² 47 C.F.R. § 79.1(e)(11)(i)(A).

⁴³ 47 C.F.R. § 79.1(e)(11)(i)(B).

⁴⁴ 47 C.F.R. § 79.1(e)(11)(i)(C).

⁴⁵ 47 C.F.R. § 79.1(e)(11)(i)(D).

⁴⁶ 47 C.F.R. § 79.1(e)(11)(i)(E).

⁴⁷ 47 C.F.R. § 79.1(e)(11)(i)(F).

out ENT if the quality of news programming did not improve, and required the preparation of a progress report by broadcasters in consultation with Consumer Groups within a year.⁴⁸

In late 2015, the National Association of Broadcasters (NAB) released a report noting mixed progress toward ENT improvements but urging the Commission to retain the availability of ENT.⁴⁹ The Consumer Groups submitted a reply to the report, accompanied by an extensive survey of local news viewers who are deaf or hard of hearing, noting persistent problems with captioning of broadcast news programming and again urging the Commission to phase out the use of ENT.⁵⁰

In a joint ex parte filing in 2016, the Consumer Groups and NAB noted that ENT was “not an ideal long-term solution for the captioning of local news programming,” with NAB noting concerns about the availability of qualified captioners and urging a path forward that would allow for the development of automated speech-to-text software.⁵¹ Consumer Groups expressed concern over the pace of development of speech-to-text technology and urged the Commission to proceed with phasing out allowing the use of ENT, but endorsed the notion of allowing automated speech-to-text technologies so long as they were accompanied by objective metrics for caption quality to facilitate objective comparisons to real-time captioning.⁵² In response, the National Court Reporters Association (NCRA) noted the wide availability of qualified human captioners to fill needs for real-time captioning outside the top 25 markets and endorsed phasing out ENT, but noted that an “abundance of factors” could lead to poor-quality live captioning in some circumstances.⁵³

Aside from hosting a May 2019 forum on issues with the captioning of local news,⁵⁴ the Commission has let this impasse stand, remaining essentially silent on how to substantively improve non-technical caption quality in the more than five years it issued the *2014 Caption Quality Order*. Notwithstanding the Commission’s commitment to revisiting the efficacy of the best practices for

⁴⁸ *2014 Caption Quality Order*, 29 FCC Rcd. at 2272, ¶ 82.

⁴⁹ Docket No. 05-231 (Oct. 27, 2015), <https://www.fcc.gov/ecfs/filing/60001305446>.

⁵⁰ Docket No. 05-231 (Nov. 12, 2015), <https://www.fcc.gov/ecfs/filing/60001309321>.

⁵¹ Docket No. 05-231 at 1-2 (Feb. 16, 2016), <https://www.fcc.gov/ecfs/filing/60001431265>.

⁵² *Id.* at 2 (Mar. 16, 2016), <https://www.fcc.gov/ecfs/filing/60001502751>.

⁵³ Reply Comment of NCRA, Docket No. 05-231 at 2.

⁵⁴ <https://www.fcc.gov/news-events/events/2019/05/forum-captioning-local-news-programs>.

live captioning, the Commission has not taken any formal action to develop a record or launched any inquiry into the ongoing quality of captions for live programming.

III. Quality problems with captions for live programming persist widely under the “best practices” regime.

Since the adoption of human- and ENT-centric “best practices,” Consumer Groups have continued to receive widespread complaints from consumers that quality problems with captions of live programming across a range of markets have continued to persist and even deteriorate in some cases. For example, an informal survey by HLAA of more than 900 of its members in May and June of 2019, a summary of which is attached as an appendix to this petition,⁵⁵ revealed that consumers across the country continue to encounter significant problems with the captioning of live programming. For example:

- **Missing Captions for Sports and Weather.** Significant gaps continue to persist in the captioning of sports and weather; with approximately twenty percent of survey respondents reporting that their local news stations do not caption weather programming and approximately a quarter reporting that the stations do not caption sports programming.⁵⁶
- **Poor/Bad Accuracy Overall Quality.** Fewer than a quarter of respondents reported that the accuracy of their local news stations’ captioning is “Good” and only sixteen percent reported that the overall quality was “Good,” with more than thirty percent responding that the accuracy is “Poor” or “Bad” and forty percent responding that the overall quality is “Poor” or “Bad.”⁵⁷
- **Missing Speaker Identification.** Only approximately ten percent of respondents report that their local news stations “Always” identify the speaker of captioned dialogue, with

⁵⁵ Appendix A contains the summary results of the survey. The Captioning DRRP and Consumer Groups are conducting detailed analysis of the textual responses and will submit that analysis and the full set of textual responses for the record when the Commission seeks public notice on this petition.

⁵⁶ See Appendix A at 7.

⁵⁷ See *id.* at 8, 14.

almost ninety percent responding that speakers are identified only “Sometimes” or “Never.”⁵⁸

- **Captions Out of Sync.** Only approximately ten percent of respondents report that their local news stations “Always” correctly synchronize the captions with dialogue, with approximately ninety percent responding that captions are only “Sometimes” or “Never” in sync with speech.⁵⁹
- **Missing Captions of Background Noises.** Only five percent of respondents report that their local news stations “Always” caption background noises such as sirens and animals, with just under fifty percent reporting that background noises are captioned only “Sometimes” and just under fifty percent reporting that background noises are “Never” captioned on their stations.⁶⁰
- **Programs Incompletely Captioned.** Fewer than a quarter of respondents report that their local news stations “Always” caption news programming completely, from beginning to end, while more than three quarters of respondents report that programs are completely captioned only “Sometimes” or “Never.”⁶¹
- **Placement Issues.** Nearly three quarters of respondents report that their local news stations “Sometimes” place captions over important information or over someone’s face.⁶²

The qualitative responses to the survey likewise describe hundreds of variations on significant issues, including many of the same problems that the Commission observed in the run-up to the *2014 Caption Quality Order*, the 2010 record refresh, the request for comments on the *2004 Petition*, and the initial development of captioning rules in the 1990s.⁶³ This survey indicates that the quality

⁵⁸ *See id.* at 9.

⁵⁹ *See id.* at 10.

⁶⁰ *See id.* at 11.

⁶¹ *See id.* at 12.

⁶² *See id.* at 13.

⁶³ As noted *supra* in footnote 55, the DRRP and Consumer Groups will submit the full set of survey responses and corresponding analysis in a future filing.

of captioning for live programming is continuing to fall short of the Commission's requirements of accuracy, synchronicity, completeness, and placement,⁶⁴ even taking into account the Commission's allowance for de minimis errors for live captioning,⁶⁵ in the absence of objective metrics.

These ongoing caption quality problems are not restricted to ENT and ASR. Even live human captioners can substantially omit or alter content to a degree that the original communicative intent of the audio track is no longer preserved. Because human captioners excel at editing captions, viewers who are deaf or hard of hearing may not know how much the captions differ from audio.

The following example illustrates this problem. This excerpt was transcribed from a live broadcast of the Oscars on February 24, 2019, in Washington DC (a top-25 market) with live human captioning. The recording was taken from ABC 7 with CEA-608 captions on cable, without detectable transmission errors, and covers the time codes where production designer Hannah Beachler was honored for her work on the movie *Black Panther* and gave her acceptance speech. The captions did not have obvious errors for those who were unable to hear the audio. However, the captions omitted about half of what Beachler actually said, and the omissions substantially altered the meaning of Beachler's speech.

The captions shown on the left are as broadcast, covering a period of 18 seconds, at a rate of 103 words per minute (WPM), substantially below the speeds at which the average viewer reads and people talk. The actual content is shown on the right, at a rate of 196 WPM:

⁶⁴ See 47 C.F.R. § 79.1(j)(2).

⁶⁵ See 47 C.F.R. § 79.1(j)

Captions:

AND I'M STRONGER BECAUSE OF THE
HART DEPARTMENT CREW, WHO
BROUGHT THEIR COURAGE AND
HUMILITY TO THE TABLE.
AND THANK YOU TO THE
WONDERFUL
CAST, WHO BREATHED LIFE INTO
THIS WORLD.

Actual Content:

And I'm stronger because of
the art department crew
who led with their hearts
and brought their experiences, talent,
courage, humility and
hard work to the table.
Supervising art director, Allan Hook
and set decorator, Jay Hart, thank you.
I am stronger today because
of this wonderful cast,
who everyday, stepped into this world
and breathed life into it.

Compared to what was actually captioned on the broadcast, the following information has been omitted or altered:

1. Beachler thanked the art director and set decorator, both by name. There is no mention of that in the original captions.
2. There is no mention of several virtues brought by the art crew: the heart in the effort, nor the talent, nor the hard work—only courage and humility.
3. There is no mention that Beachler is stronger because of the cast—which is a different intent and nuance compared to thanking someone. In fact, it appears that the thank you directed at the art crew supervisors was misplaced and applied to the cast.

IV. The human and ENT-centric, quality “best practices” approach provides no means for objectively assessing how increasingly diverse captioning methodologies compare and whether they result in accessible video programming.

Since the Commission adopted the human- and ENT-centric “best practices” approach in the *2014 Caption Quality Order*, the provision of live captions has continued to evolve and unfold with little meaningful oversight from the Commission. In addition to real-time human captioners and ENT, automatic speech recognition (ASR) captioning technologies have now begun to proliferate widely. One vendor of ASR technologies, AppTek, briefed the Commission in 2016 about then-state-of-the-art ASR captioning solutions specifically designed for broadcast news use, contending

that its ASR technology could significantly improve the accuracy and latency of news captions.⁶⁶ Since then, the Consumer Groups have received numerous reports of broadcast stations, both inside and outside the top 25 markets, migrating from both live human captioners and ENT captioning to ASR solutions of varying provenance. As with human captioning and ENT, these ASR solutions vary widely in terms of various dimensions of quality, including accuracy and completeness.

As a result, the Commission now faces an increasingly diverse landscape of captioning methodologies and technologies that are frequently yielding poor-quality captions for millions of consumers. From our meetings over the past five years with NAB, the National Cable and Telecommunications Association, caption vendors, and ASR technology providers and viewing a variety of captioning examples across a variety of videos, it is clear that quality problems are not restricted to one methodology or technology. Each technology and methodology—human, ENT, ASR, and hybrid models—demonstrates promise in some contexts but suffers from quality problems in others, and the “best practices” approach provides no means for assessing or addressing the tradeoffs between them, leaving the primary driver as the cost, not the quality, of the captions.

V. The Commission should launch an inquiry to develop objective metrics for caption quality and processes for enforcing them.

While we remain committed to ongoing dialogue with members of the captioning, broadcast, and cable industries, it is becoming clear against the increasingly diverse backdrop of captioning technologies and methodologies that methodology-specific, quality-agnostic “best practices” are not up to the task of vindicating the civil rights of Americans who are deaf or hard of hearing to access live video programming on equal terms. To avoid a race to the bottom on cost, the Commission should adopt a more rigorous substantive and procedural framework for assessing the quality of

⁶⁶ *Ex Parte of AppTek*, Docket No. 05-231 (July 29, 2016), <https://www.fcc.gov/ecfs/filing/10729264643292>.

captions in technology-neutral terms that focuses on the actual consumer experience rather than substituting methodology as a proxy.

Dating back to the *1997 Order*, the Commission has declined to delve into the topic of metrics in part on the grounds that doing so is too difficult.⁶⁷ We acknowledge that this task has not become substantially simpler in the intervening time as the understanding of the relationship between dimensions of captioning quality and the accessibility of video programming have evolved.

Nevertheless, a new research effort, the Disability and Rehabilitation Research Project on Twenty-First Century Captioning Technology, Metrics and Usability, conducted by the Captioning DRRP, is poised to provide the basic research to answer some of these questions.⁶⁸ Over the coming years, the project will develop rigorous, scientifically sound, consumer-focused metrics for captioning quality and accompanying methods to conduct aggregate quality evaluations across the video programming ecosystem.⁶⁹

To incorporate the results of the project, the Commission should immediately initiate an inquiry into the state of the art of closed captioning techniques for live television programming and how the varying dimensions of caption quality, including accuracy, synchronicity, completeness, and placement affect the accessibility of video programming. In addition to the DRRP, the inquiry will afford all stakeholders, including Consumer Groups, video programmers, captioning vendors, technology developers, and members of the public to provide input on metrics for captions.⁷⁰

Following the development of a record in response to that inquiry, the Commission should use that record to adopt technology- and methodology-neutral metrics and use them to set objective thresholds that video programmers must satisfy to ensure that the live programming they deliver is

⁶⁷ *1997 ReO*, 13 FCC Rcd. at 3374, ¶ 222.

⁶⁸ *Twenty-First Century Captioning Technology, Metrics and Usability*, <https://captions.us/cms/> (last visited July 5, 2019).

⁶⁹ *See id.*

⁷⁰ While we believe the primary focus and intended output of the proposed inquiry should be captioning metrics, we would welcome the inclusion of discussion on the economics of closed captioning techniques and methodologies.

in fact accessible to Americans through the provision of high-quality captions, regardless of how those captions are generated.

The Commission's outmoded practice of relying on consumer complaints does not work for caption quality, where consumers who rely exclusively or primarily on captions to convey the content of video programming often are not in a position to perceive problems with the content of captions. Accordingly, the Commission should adopt a rigorous monitoring and compliance regime, routinely sampling and evaluating the quality of captions for live programming, to ensure that Americans who are deaf or hard of hearing need not bear the sole responsibility or the costs of verifying the quality of captions.

VI. The Commission must provide immediate guidance to video programmers on the permissible use of automatic speech recognition.

Compiling a record on metrics will take a significant investment of time on the part of the Commission and stakeholders. In the meantime, it is incumbent on the Commission to ensure that inexpensive but poor-quality ASR techniques do not permanently undermine the market for live human captioners with no guarantee that ASR will not perpetuate or exacerbate the quality problems described above. At the very least, the Commission must ensure that ASR lives up to the current level of quality facilitated by the best practices for human captioners.

Unfortunately, it is unclear how the use of ASR techniques fit into the existing best practices, which at least implicitly contemplate human captioners in a number of respects. For example:

- The best practices for video programmers include training provisions that apply to “employees and contractors who provide caption services” and require “ensur[ing] that there is oversight of individual captioners’ performance.”⁷¹ The video programmer best practices also require providing “advance access to preparation materials.”⁷²

⁷¹ 47 C.F.R. § 79.1(k)(1)(i)(c).

⁷² 47 C.F.R. § 79.1(k)(1)(ii)(A).

- The best practices for real-time captioning vendors require ensuring “proper screening, training, supervision, and evaluation of captioners by experienced and qualified real-time captioning experts” and “that captioners are qualified for the type and difficulty level of the programs to which they are assigned.”⁷³
- The best practices for real-time captioners require captioners to “[p]repare as thoroughly as possible for each program,” “[f]ile thorough discrepancy reports with the caption vendor in a timely manner,” “[m]onitor captions to allow for immediate correction of errors and prevention of similar errors appearing or repeating in captions,” “[p]erform frequent and regular self-evaluations,” “[p]erform regular dictionary maintenance,” “[k]eep captioning equipment in good working order and update software and equipment as needed,” “[p]ossess the technical skills to troubleshoot technical issues,” and “[k]eep abreast of current events and topics that they caption.”⁷⁴

It is unclear to us how users of ASR technologies purport to comport with these best practices. For example, how does ASR technology perform “self-evaluations” or “possess” the technical skills to troubleshoot technical issues? How does a caption vendor properly screen, train, and supervise an ASR system? How does a video programmer ensure the oversight of an ASR system?

Accordingly, following a public notice soliciting comment on this petition, the Commission should promptly issue a declaratory ruling pursuant to Rule 1.403,⁷⁵ or alternatively an expedited order under Rule 1.412(c) adopting changed rules,⁷⁶ explaining how the Commission’s “best practices” for video programmers, caption vendors, and live captioners apply to the use of ASR technologies by programmers, vendors, or captioners.

It is not clear to us whether the existing human-centric best practices are sufficiently amenable to application by analogy to ASR for the Commission to issue a declaratory ruling, or whether the

⁷³ 47 C.F.R. § 79.1(k)(2)(viii), (x).

⁷⁴ 47 C.F.R. § 79.1(k)(3)(v), (vi), (ix)-(xiv)

⁷⁵ 47 C.F.R. § 1.403.

⁷⁶ 47 C.F.R. § 1.412(c).

application of the best practices to ASR would represent such a significant departure that changed rules are necessary—a question on which the Commission should solicit immediate comment in placing this petition on public notice. Given the urgent nature of this issue, if the Commission determines that changed rules are necessary, it should act with haste and issue expedited changed rules under Rule 1.412(c) based on these initial comments without issuing a formal notice of proposed rulemaking and undergoing an additional comment cycle.⁷⁷

If the Commission cannot quickly resolve the applicability of the real-time captioning best practices to ASR through the issuance of a declaratory ruling or expedited rule change, it must promptly issue a declaratory ruling that video programmers using ASR cannot comply with Rule 79.1(m)(1)(ii) through the use of best practices. The Commission should further clarify that compliance with Rule 79.1(m)(1)(ii) requires video programmers to either cease the use of ASR and contract with captioning vendors using human captioners consistent with the best practices, or instead certify in detail to the Commission that their ongoing use of ASR comports with the quality standards in Rule 79.1(j)(2). Finally, the Commission should clarify that to certify compliance with the quality standards under Rule 79.1(j)(2) using ASR technologies, video programmers must investigate, verify, and explain how the use of ASR technology in the supply chain for captions used on their live programs results in captions that satisfy the accuracy, synchronicity, completeness, and placement requirements of Rule 79.1(j)(2).

⁷⁷ Pursuant to Rule 1.412(c), the Commission may waive the public notice provisions of the Administrative Procedure Act when doing so is contrary to the public interest. *See, e.g., Establishment of Policies & Serv. Rules for the Broad.-Satellite Serv.*, 22 FCC Rcd. 8842, 8898, ¶ 138 (2007)