MachineGenius

May 30, 2018

via ECFS and e-mail

Erik Strand
President, MachineGenius, Inc.
52 Dean Rd
Holliston, MA 01746
erik.strand@machinegenius.com

Marlene H. Dortch
Secretary
Federal Communications Commission
445 12th Street, SW, Room TW-A325
Washington, DC 20554

Re: ***Telecommunications Relay Services and Speech-to-Speech Services for Individuals with Hearing and Speech Disabilities • CG Docket No. 03-123 Misuse of Internet Protocol (IP) Captioned Telephone Service • CG Docket No. 13-24***

Dear Ms. Dortch,

In response to the May 17, 2018 draft Report and Order, Declaratory Ruling and Further NPRM[1] in the above-referenced proceedings and the subsequent May 25, 2018 joint *ex parte* filing[2] submitted by Telecommunications for the Deaf and Hard of Hearing, Inc. (TDI), Hearing Loss Association of America (HLAA), and Gallaudet University Technology Access Program (TAP), MachineGenius seeks to clarify its position on aspects of the use of fully-automated Automated Speech Recognition (ASR) for delivering IP CTS service.

MachineGenius joins TDI, HLAA, and Gallaudet University TAP in supporting the Commission's efforts to ensure the sustainability of the IP CTS program and to shift to a technology-neutral framework for evaluating IP CTS providers. However, while we highly value the input of these IP CTS stakeholders, we believe many of their expressed concerns with the draft's Declaratory Ruling on the use of ASR technology are poorly substantiated. We address these point by point below with excerpts from the TDI filing in italics.

The Declaratory Ruling's approach of delegating to the Bureau the responsibility of applying the existing TRS minimum standards to IP CTS applicants proposing to use ASR does not

---

[1] Draft Report and Order, Declaratory Ruling, Further Notice of Proposed Rulemaking, and Notice of Inquiry (May 17) ("Draft Item"), https://www.fcc.gov/document/reforming-internet-protocol-captioned-telephone-service.
[2] Ex Parte of Telecommunications for the Deaf and Hard of Hearing, Inc. (TDI), Hearing Loss Association of America (HLAA), Gallaudet University Technology Access Program (TAP) (May 25, 2018) ("TDI Filing"), https://ecfsapi.fcc.gov/file/105252245210428/2018.05.25%20IP%20CTS%20Draft%20Item%20Ex%20Parte%20final.pdf.

acknowledge that the minimum standards are replete with explicit references to human communications assistants (CAs).[3]

On the contrary, the Declaratory Ruling explicitly and extensively addresses the role of communications assistants (CAs), and the fact that the CA role may be limited or nonexistent in an ASR-based IP CTS solution[4], concluding that "the definition of IP CTS in our rules does not specify how captions must be generated, including whether they should be generated through automation or human assisted methods." It observes that "use of [ASR] technology for IP CTS without CA involvement does not fundamentally change the functional role of the service, which is to produce captions from a user's speech.[5]" It is in this context that the Declaratory Ruling authorizes the approval of ASR-based IP CTS applications, provided an applicant meets current minimum mandatory standards as expressed in the existing rules, except as otherwise waived.[6]

MachineGenius maintains that references to CAs in the Commission's rules do not preclude the body of existing rules from applying to an ASR-based IP CTS provider, and do not weaken the body of rules in the enforcement of minimum standards for ASR-based providers.

The approach taken by MachineGenius to satisfy the minimum-standards rules can be applied generally across the class of ASR-based solutions:

- To the extent that a minimum standard happens to be CA-specific in a way that by its nature is inapplicable to ASR-based solutions, MachineGenius has requested a waiver of such rule(s). For example, MachineGenius has requested a waiver of the CA typing-speed rule, and of the requirement to provide a CA identification number in reports to the Administrator.

- Where, despite the mention of CAs in a rule, the underlying purpose of that rule remains applicable to ASR, we have sought a waiver only "to the extent that the rule applies specifically to CAs", and not sought a waiver from meeting the minimum standard that is the underlying purpose of the rule. In all cases where MachineGenius has requested a waiver on these grounds, the MachineGenius solution meets or exceeds the performance requirements that are the underlying purpose of the rule. For example, MachineGenius has sought a waiver of the confidentiality and conversation content rule only to the extent that it applies to CAs specifically, and remains committed to meeting the confidentiality minimum standards that are the underlying purpose of the rule.

- Where a rule does not relate to CAs, MachineGenius has not requested a waiver and is committed to meeting the minimum standard verbatim.

---

[3] *Id.* at 2.
[4] Draft Item, §§ B,C
[5] *Id.* ¶ 53.
[6] *Id.* ¶ 58.

This approach of common-sense application of waivers is consonant with the provisions of the Declaratory Ruling. The complete MachineGenius waiver request[7] can be found as an attachment to our IP CTS application.[8]

The Declaratory Ruling's approach of delegating to the Bureau the responsibility of applying the existing TRS minimum standards to IP CTS applicants proposing to use ASR … provides little guidance as to how the Bureau should evaluate compliance with those standards by machine-learning algorithms.[9]

It is unclear why further guidance in applying the existing TRS minimum standards is necessary. Current IP CTS providers already leverage "re-voiced" ASR; fully-automated ASR differs only by removing the necessity for re-voicing; the current rules are not so complex or CA-dependent that they cannot be applied to a fully-automated ASR-based system. As detailed above, cases where the rules apply specifically and exclusively to CAs, and have no application to an ASR-based system, are obvious and can be waived. Rules where CAs are mentioned but the underlying purpose of the rule remains applicable to an ASR-based system can and should remain in force and be waived only to the extent that they are CA-specific. Rules that are independent of the method of generating captions remain in force in their entirety.

Further, "machine-learning algorithms" as they relate to ASR only play a role in association with the accuracy and latency of captions. Speed-to-answer, system availability, dropped call rates, reporting requirements, and all other minimum-standards requirements are features of distinct parts of the system unrelated to caption accuracy and latency, and are all quantifiable without reference to the method of generating captions.

Given the above, and assuming the comparative latency advantages when using ASR are undisputed and uncontroversial, we assume that the concern expressed in the TDI filing pertains specifically to accuracy. There is only one minimum-standard rule (*viz.* 64.604(a)(2)(ii)) that addresses accuracy, and the underlying purpose of the rule is straightforwardly applicable to ASR-based solutions. Any further guidance necessary to evaluate compliance with this rule by ASR-based solutions would apply equally to CA-based solutions. The same evaluative measures can and should be used in either case.

MachineGenius submits that the absence of further guidance specific to ASR does not represent a deficiency in the Declaratory Ruling.

The Declaratory Ruling raises significant legal issues and [we] are concerned that this approach will result in the Bureau issuing determinations that effectively result in the modification of the minimum standards without solicitation of public comment, potentially giving

---

[7] MachineGenius, Inc. Request for Waiver (October 13, 2017), https://ecfsapi.fcc.gov/file/1014215719459/IPCTS%20Waiver%20Request_PUBLIC.pdf.
[8] MachineGenius Application (October 13, 2017), https://ecfsapi.fcc.gov/file/1014215719459/IPCTS%20Application_-%20PUBLIC%20NON-CONFIDENTIAL.pdf.
[9] TDI Filing at 2.

rise to violation of the notice and comment requirements of Administrative Procedure Act and miring the deployment of IP CTS improvements in litigation.[10]

As detailed above, application of the current rules to ASR does not "effectively result in the modification of the minimum standards" so long as common-sense waivers of inapplicable provisions of the minimum-standards rules are granted.

The Declaratory Ruling opens the door for ASR solutions to widespread deployment without the implementation of quality standards or performance metrics. While the Commission acknowledges that Section 225 of the Communications Act requires it ensure that IP CTS solutions provide functional equivalence to consumers with disabilities, the draft item relegates this task to a Notice of Inquiry with no obvious timeline while immediately opening the door to the deployment of ASR solutions with potentially serious quality shortcomings.[11]

MachineGenius considers this to be an extremely unlikely scenario. It is unclear what the incentive would be for an existing IP CTS provider to deploy an ASR solution with serious quality shortcomings; it would be immediately rejected by consumers. Likewise, as a practical matter, any existing IP CTS provider would not engage in widespread deployment of an ASR solution without first testing it for quality on a small scale.

With respect to prospective IP CTS providers like MachineGenius, given the intensive application evaluation process for IP CTS providers (ASR-based or otherwise), and the provision for case-by-case review and approval documented in the Declaratory Ruling[12], it is unclear how any solution with serious quality shortcomings could reach approval. Similar to existing providers, any prospective provider is strongly disincentivized to seek approval for a solution with serious quality shortcomings, as it will be quickly rejected by consumers.

Conversely, it is an extremely likely scenario that existing or prospective IP CTS providers will choose (as MachineGenius has) a state-of-the-art ASR engine that has a track-record of proven performance and is demonstrably free of serious quality shortcomings.

These quality concerns are not mere speculation. For example, the Commission has already received IP CTS applications contemplating the use of ASR technology from MachineGenius *and* VTCSecure *that do not adequately address the quality of the underlying technology … MachineGenius's public application is replete with vague promises of "high-quality" and lower-cost captions, but requests waivers of numerous minimum standards.[13]*

MachineGenius uses the term "high-quality captions" in a straightforward way to refer to high-accuracy, high-readability, high-comprehensibility captions. We further detail these concepts in

---

[10] *Ibid.*

[11] *Ibid.*

[12] Draft Item, ¶ 60, declaring "We do not, at this time, prescribe the specific manner in which a provider must use ASR in order to be certified. Rather, in reviewing specific certification applications, the Bureau may determine on a case-by-case basis the extent to which an applicant's proposed method of providing ASR will enable it to provide IP CTS in a manner that meets the Commission's minimum TRS standards for functionally equivalent service."

[13] TDI Filing at 2.

the Automatic Speech Recognition for IP CTS exhibit filed under seal with our Application[14], and presume that these concepts make intuitive sense even absent review of that proprietary document.

While the above statement of concern by TDI strongly implies that MachineGenius has requested waivers intended to exempt our solution from providing high-quality captions, that is incorrect. As stated above, we have only requested waivers for rules and portions thereof that are inapplicable to ASR-based solution, and we have not requested the waiver of any standard that pertains to caption quality.

We assume that the claim of substantially lower costs associated with ASR-generated captions is uncontested.

MachineGenius' claims of high-quality captions are not vague promises; we have demonstrated the performance of our solution in live, unscripted demos to FCC stakeholders, and voluntarily made preview releases of our application available to MITRE for independent evaluation. We have also cited concrete performance measures published by major ASR providers, including MachineGenius' ASR provider, and provided an exhibit with our application that presents factual evidence for the efficacy of state-of-the-art ASR as an IP CTS solution.[15]

We are also concerned that the Declaratory Ruling leaves open serious questions about protecting the privacy of sensitive conversations conducted over IP CTS systems. For example, the Declaratory Ruling declares that conversations must be "kept confidential" but appears to contemplate that ASR providers can use internet-based ASR engine providers, which inherently require transferring call recordings to third-party providers.[16]

The inference that we are invited to draw is that using a 3rd-party to generate captions precludes keeping a conversation confidential. This is plainly false. If it were true, then current providers utilizing outsourced CAs would also be unable to keep conversations confidential, and this has not been judged to be the case. It is also not the case that "call recordings" are transferred to 3rd-party ASR providers; only live, real-time, audio is sent, and there is no confidential "data at rest" on the servers of either an IP CTS provider or a 3rd-party ASR provider in a properly-designed ASR-based IP CTS solution.

On the contrary, calls handled by ASR are in fact more private than calls handled by CAs. First, by definition, it is definitionally true that calls handled by CAs have at least one 3rd-party human listening to them, while calls handled by ASR do not. Second, it is also extremely likely that the physical, logical, and network security afforded by the "battle tested" datacenters of the major ASR providers exceeds the level of security afforded by CA-staffed contact centers, and their local networks and application infrastructure.

---

[14] MachineGenius Application, Exhibit A, Automatic Speech Recognition for IP CTS (filed under seal) (October 13, 2017).

[15] *Ibid.*

[16] TDI Filing at 3.

Additionally, all data sent to and from a cloud-based provider in the MachineGenius solution is encrypted with TLS, and this would presumably be the case for any ASR-based IP CTS provider using a cloud-based ASR service. Therefore, neither the *per se* act of transferring data, nor the processing of that data by a major cloud-based ASR vendor, represents a particular challenge (versus CA-based systems) to keeping calls confidential.

Many engine providers collect and utilize audio recordings for the purpose of improving the accuracy of the underlying machine learning technology that powers their engines. The Declaratory Ruling leaves unclear the Commission's intent for how the Bureau should approach the complex tradeoffs between privacy and quality improvements inherent in the use of machine learning algorithms for voice transcription, the Commission's expectations for and approach to ensuring that the use of third-party voice engine providers does not result in the unlawful disclosure of sensitive call information, or whether and how particular practices and interactions between ASR-based IP CTS providers and ASR engine providers will comply with relevant legal regimes such as the Wiretap Act.[17]

It is true that many ASR providers can utilize submitted audio to improve the accuracy of their engine. However, it is also true that each of the major engine providers offers an operational mode where no audio or transcript data is collected for any purpose whatsoever, and in which all data disappears immediately after speech recognition is performed, rendering moot concerns related to data collection.

It is also not necessarily the case that there is an inherent tradeoff between privacy and quality improvements, and where it may be the case, the tradeoffs are not complex.

- Even when data is retained for improving speech recognition models, the data is anonymized and fragmentary and remains subject to the privacy policy of the ASR provider;
- The quality of service of the major providers will continue to improve independently of whether or not IP CTS call audio is utilized for the purposes of improving ASR service;
- If it is concluded that collecting data for service improvement represents a security risk, and it is also concluded that ASR output will improve if data collection is enabled, then the straightforward tradeoff is a minimal but potentially acceptable decrease in privacy in return for a modest increment in caption quality. Again, this is conditioned upon assumptions, and if such a tradeoff is undesirable, or if the Bureau requires that no IP CTS audio be used for the improvement of ASR service (in order to comply TRS call confidentiality rules[18]) then modes of ASR processing where no data is collected can be utilized.

With respect to compliance with the Wiretap Act and related laws, it is difficult to see how they would apply any differently to ASR-based IP CTS than to CA-based IP CTS, both of which may

---

[17] *Ibid.*

[18] 47 CFR § 64.604(a)(2).

leverage 3rd-party providers, and differ substantively only in the method used to generate captions.

As with quality, existing applications before the Commission fail to sufficiently address these concerns. For example, the VTCSecure public filing does not appear to specifically address confidentiality for the ASR portion of its proposed service. And the MachineGenius public filing vaguely states that "[a]udio and transcripts of calls are not stored remotely" but notes that the service is subject to unspecified privacy policies, including of an unidentified third-party engine provider.[19]

By "audio and transcripts of calls are not stored remotely" MachineGenius means that these data will not be stored on our servers (i.e. remotely), but may be stored (especially captions) locally to the user's client device for purposes of display to the user. This is the same operational behavior of current IP CTS providers, and our statement is intended to confirm compliance with the minimum standards for data retention.[20]

The privacy policies of all the major cloud-based ASR providers are publicly available; MachineGenius makes clear in its Application that these privacy policies will pertain to the handling of IP CTS data. This is the direct analog of a CA-based IP CTS provider confirming that IP CTS data will be subject to the privacy policies of an outsourced CA vendor. In both cases, the privacy policy in question must be sufficiently strong to permit the IP CTS provider to meet relevant mandatory minimum standards. MachineGenius' ASR provider is identified in redacted portions of our application.

As noted above, independent of any privacy policy, performing speech recognition in a mode that collects no data is possible with all of the major ASR providers, and any risk associated with transfer of data to a 3rd-party for the purpose of generating captions applies equally to ASR- and CA-based providers.

Finally, we agree with Hamilton Relay that the Declaratory Ruling does not sufficiently address the interaction of ASR-based IP CTS providers with 9-1-1 to ensure their ability to safely handle emergency calls. To forge ahead with ASR-based solutions without confidence that they will work properly in an emergency could seriously jeopardize the lives and safety of consumers with disabilities.

MachineGenius agrees that both the Bureau and consumers need to be confident that emergency calls will be safely handled. To that end, MachineGenius will undertake a study to show the performance of our solution on 9-1-1 calls. If the results of that study reveal that the performance is not at least equal to the performance of CA-based IP CTS solutions on similar calls, then we will commit to staffing CAs for the handling of emergency calls[21].

---

[19] TDI Filing at 4.

[20] 47 CFR § 64.604(a)(2).

[21] The MachineGenius ASR-based system is designed to work primarily in fully-automated mode, but can also be used in "revoicing mode" where CAs can listen to live call audio and input or correct captions by speaking or typing. This functionality will enable us to support CAs for 9-1-1 calls, if deemed necessary.

While we believe these concerns can be overcome with the development of a detailed record and further dialogue among stakeholders in this proceeding, the Commission risks substantial harm to the civil rights of consumers who are hard of hearing by proceeding with the deployment of ASR technologies without developing a more rigorous approach to ensuring quality and privacy.

... [We urge] at a bare minimum, to put out on public notice and solicit public comment on all IP CTS applications, specifically on quality and privacy issues, to ensure that quality and privacy issues are not overlooked when the Bureau considers ASR (and other) IP CTS applications, and insert specific language in the Declaratory Ruling making clear that the Commission will not approve ASR-based IP CTS services that do not deliver functionally equivalent quality and protect the privacy and confidentiality of consumers with disabilities.[22]

It is unclear how permitting the deployment of ASR technologies might lead to a substantial harm to the civil rights of consumers, especially given that the Commission is not mandating ASR as the sole means of offering IP CTS, and that in the near-term ASR-based IP CTS will remain a nascent service.[23] In contrast, for the reasons well-documented in the Declaratory Ruling, MachineGenius submits that it will certainly be contrary to the public good – both for hard-of-hearing consumers and for the viability of the TRS Fund – to further delay Commission action to approve ASR-based IP CTS solutions.

As for the recommendation to put IP CTS applications on public notice, and to insert specific related language into the Declaratory Ruling, MachineGenius is confident that it is already currently the case that the Bureau will not overlook quality and privacy issues when considering ASR-based IP CTS applications, and that the Commission will not approve IP CTS services that do not deliver functionally equivalent quality and protect the privacy and confidentiality of consumers with disabilities.
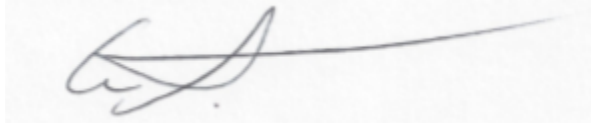
MachineGenius urges the Commission to:

- Approve the Declaratory Ruling in the draft item;
- Require IP CTS stakeholders to swiftly resolve the issues in the draft item's Notice of Inquiry, including the formulation of IP CTS performance goals and service quality measures that apply equally to all providers of IP CTS, regardless of underlying delivery technology;
- Reconsider the draft item's approach to eligibility certifications, which we (along with TDI, HLAA, and Gallaudet University TAP) believe will disenfranchise consumers from acquiring access to necessary IP CTS while adding unnecessary complexity and cost.[24]

---

[22] TDI Filing at 4.
[23] Draft Item, ¶ 60.
[24] TDI Filing at 5.

Respectfully submitted,

Erik Strand
President
MachineGenius, Inc.
erik.strand@machinegenius.com

CC:
Nirali Patel, Office of Chairman Pai
Zenji Nakazawa, Office of Chairman Pai
Karen Peltz Strauss, Consumer and Governmental Affairs Bureau
Robert Aldrich, Consumer and Governmental Affairs Bureau
Michael Scott, Consumer and Governmental Affairs Bureau
Eliot Greenwald, Consumer and Governmental Affairs Bureau
David (Raster) Schmidt, Office of the Managing Director
Andrew Mulitz, Office of the Managing Director