

Contact Center AI : Google muscle la reconnaissance vocale

Google a mis à niveau et enrichi sa suite de solutions de reconnaissance vocale basées sur l'intelligence artificielle pour centres d'appels.

L'ensemble est regroupé au sein du portefeuille [Contact Center AI](#) depuis l'été 2018.

Le service est orienté développeurs et [chatbots](#). Les améliorations et mises à jour annoncées concernent les technologies socles de l'offre.

Ainsi, Dialogflow, la plateforme de développement d'[interfaces conversationnelles](#) de Google, inclut désormais la fonctionnalité [Auto Speech Adaptation](#) (en version bêta).

Celle-ci peut être activée par le développeur dans la console Dialogflow.

Avec elle, les agents virtuels devraient pouvoir « comprendre rapidement ce dont les clients ont besoin », y compris dans des environnements bruyants, a indiqué Google dans un [billet de blog](#). La fonctionnalité améliorerait la précision de ces bots de plus « de 40% ».

SpeechContext

D'autres améliorations, toutes en version bêta, concernent Cloud Speech-to-Text, l'API de Google qui permet aux développeurs de convertir du son en texte en exploitant des modèles de réseaux de neurones. La solution intègre dorénavant un modèle de base de reconnaissance vocale amélioré pour les solutions de serveur vocal interactif (Interactive Voice Response – IVR) et les agents virtuels « phone-based ».

La solution inclut également une fonction étendue de streaming (Endless streaming). « Jusqu'à présent, Cloud Speech-to-Text prenait en charge la diffusion audio uniquement par sessions d'une minute. Ce qui posait problème pour les transcriptions de longue durée comme les réunions, la vidéo en direct et les appels téléphoniques. » C'est la raison pour laquelle cette durée « a été portée à 5 minutes », a souligné Google.

Par ailleurs, le support natif du format de fichier MP3 est activé dans Cloud Speech-to-Text.

Enfin, pour Dialogflow et Cloud Speech-to-Text, des mises à jour des paramètres de SpeechContext de Google sont proposés. Avec elles, les développeurs doivent pouvoir compléter leurs applications d'informations contextuelles supplémentaires permettant d'améliorer la précision de la transcription. Et, par extension, [l'expérience client](#).