

Première mondiale : avec Kyutai, Niel, Saadé et Schmidt donnent naissance à Moshi

Moshi, l'intelligence artificielle vocale développée par le laboratoire Kyutai, est une première mondiale. Capable de reconnaître jusqu'à 70 émotions, les discussions avec cette IA générative sont d'une fluidité prodigieuse.

Temps de lecture : minute

3 juillet 2024

«Hey what'up ? What's your name ? My name is Moshi, how can I help you today ?» Moshi est née du laboratoire Kyutai, lancé il y a six mois par Xavier Niel, Eric Schmidt et Rodolphe Saadé. L'équipe de 8 chercheurs a présenté sa première innovation : une intelligence artificielle multimodale et générative... vocale. Une première mondiale. «Cela a été six mois fous !», annonce dès les premières phrases de son discours Patrick Pérez, directeur et cofondateur du laboratoire Kyutai.

Les premières démonstrations de l'intelligence artificielle Moshi sont déjà impressionnantes. Le postulat de départ de l'équipe est simple : la communication para verbale est aussi importante que la communication verbale. Le texte est la meilleure transcription d'un message oral mais il y a, tout de même, cette perte d'information apportée par le ton de la voix et les émotions. C'est pour pallier cette perte que Kyutai a créé Moshi, basée sur un modèle léger de 7 milliards de paramètres.

Moshi rend une discussion la plus fluide

Moshi peut écouter, répondre et intervenir à tout moment dans une

conversation. Elle est aussi capable de jouer un rôle et de raconter une histoire avec la voix d'un personnage. Moshi peut s'adapter à un accent, chuchoter et exprimer (et reconnaître) jusqu'à 70 émotions différentes.

Pour nourrir leur modèle d'IA générative, les chercheurs de Kyutai ont utilisé des ressources textuelles et audios à travers une base de données d'appels téléphoniques enregistrés dans les années 1990.

Pour l'instant Kyutai a développé son IA en anglais. Il lui ont enregistré la voix d'Alice, une comédienne, qui s'est prêté au jeu pendant une vingtaine d'heures. *«Une demie-heure aurait suffi»*, glisse Neil Zeghidour, chercheur membre de l'équipe depuis la création du labo.

Le code bientôt partagé en open source

Kyutai est un laboratoire open-source. Le code et les poids des modèles seront bientôt partagés librement et gratuitement. *«Nous serons ravis de discuter avec des équipes qui veulent se saisir de Moshi et la commercialiser»*, annonce Patrick Pérez. *«La communauté pourra notamment étendre la base de connaissances et la factualité de Moshi, aujourd'hui volontairement limitées.»*

Kyutai a bénéficié d'un financement de 300 millions d'euros lors de son démarrage, il y a un peu plus de six mois. *«La majorité des financements ont été utilisés pour les puissances de calcul et les GPU's»*, a confié Patrick Pérez. Pour développer Moshi, un millier de puces sont nécessaires. Kyutai a notamment noué des partenariats avec Scaleway et Hugging Face.

D'autres devraient être annoncés dans les prochains mois, avec des médias par exemple. Et pour cela, le prototype de test est déjà accessible à tous.



À lire aussi

Xavier Niel : «Je n'ai pas envie que l'on ait des modèles d'IA entraînés en dehors de l'Europe»



MADDYNEWS

La newsletter qu'il vous faut pour ne rien rater de l'actualité des startups françaises !

JE M'INSCRIS

Article écrit par Aurélie Pasquier