

ai-PULSE : Kyutai et H à l'offensive face à OpenAI depuis Station F

Pour la deuxième année consécutive, Scaleway a mis à l'honneur des acteurs clés de l'IA à Station F, à l'occasion de sa conférence ai-PULSE. Une opportunité pour le laboratoire Kyutai et la startup H de préciser leurs ambitions.

Temps de lecture : minute

8 novembre 2024

C'était l'effervescence des grands jours ce jeudi matin à Station F. Comme l'an passé, Scaleway, la filiale cloud d'Iliad, a choisi le campus parisien de startups pour tenir sa conférence ai-PULSE. Et comme en 2023, Xavier Niel, le fondateur d'Iliad et de Station F, était de la partie, comme pour mieux marquer l'importance d'un tel événement. Mais à la différence de l'édition inaugurale d'ai-PULSE, l'entrepreneur le plus connu de France ne s'est pas éternisé sur scène. C'est ensuite Aude Durand, bras droit de Xavier Niel qui est aujourd'hui la n°2 du groupe Iliad, qui a pris le relai pour animer l'ensemble de la matinée.

Dans le cadre de celle-ci, Damien Lucas, le directeur général de Scaleway, a d'abord pris la parole pour présenter les dernières avancées de son groupe pour soutenir l'écosystème IA. Le dirigeant a ainsi assuré que Scaleway disposait de la plus grande capacité de calcul dédiée à l'IA en Europe, celle-ci ayant été multipliée par cinq en un an grâce notamment aux puces GPU de dernière génération de Nvidia et d'AMD. Comme cette puissance de calcul soulève des préoccupations sur son impact sur la planète, l'entreprise a profité d'ai-PULSE pour dévoiler son calculateur d'empreinte carbone.

Scaleway au service des développeurs

Pour améliorer le quotidien des développeurs de l'écosystème IA, Scaleway a également annoncé des améliorations au niveau de son approche open source. Dans ce cadre, la filiale cloud d'Iliad a fait savoir qu'elle voulait simplifier l'accès à des modèles ouverts sans avoir à modifier une seule ligne de code.

Elle veut aussi embrasser pleinement la révolution de l'informatique quantique avec l'aide de son partenaire Quandela. Et après avoir proposé un service d'épilation quantique avec la startup tricolore, Scaleway a décidé de pousser le curseur un cran plus loin pour permettre aux développeurs d'exécuter leur code sur un véritable ordinateur quantique.

Cerise sur le gâteau, Damien Lucas a annoncé le lancement de Moshi-as-a-Service, une solution pour permettre aux développeurs de construire des premiers cas d'usage souverains avec l'IA vocale développée par Kyutai, le laboratoire de recherche open source lancé l'an passé en grande pompe lors de la première édition d'ai-PULSE.



Kyutai peaufine Moshi

Un an plus tard, Patrick Pérez, CEO de Kyutai, est d'ailleurs revenu sur la scène de Station F ce jeudi matin pour détailler l'approche de son laboratoire avec son intelligence artificielle multimodale et générative vocale, dévoilée en juillet dernier. Dans ce cadre, le chercheur a d'ailleurs indiqué que des informations clés sur le développement de Moshi, comme le code d'entraînement, allaient prochainement être publiées. «*Pour chacun de nos projets, on partage le code et nos recherches*», a souligné Patrick Pérez.

Cinq mois après le lancement de Moshi, ce sont plus d'un demi-million de sessions qui ont été lancées sur le site de démo de l'IA vocale, a assuré le patron de Kyutai. Devant l'intérêt de la communauté IA pour sa solution, l'équipe du laboratoire prévoit notamment d'étoffer les capacités linguistiques de Moshi pour que celui-ci ne soit pas limité à l'anglais.

Depuis Paris, Patrick Pérez a donc annoncé que le français serait la prochaine langue intégrée à Moshi. Passé notamment par Microsoft et l'Inria, le chercheur tricolore a également démontré l'étendue des cas d'usage possibles avec Moshi, comme le fait de pouvoir interagir avec des images, faire de la transcription automatique, ou la capacité à générer automatiquement la suite d'un fichier audio.

Si le potentiel de Moshi est prometteur, on notera tout de même une légère fausse note avec une démo en direct de Moshi avec Xavier Niel à l'entame de la conférence qui a été quelque peu laborieuse. Mais on en tiendra pas forcément rigueur, pareille péripétie s'étant déjà produite en pleine présentation de l'assistant conversationnel Bard de Google.

H se dévoile partiellement

Autre acteur français important dans l'écosystème IA, H était également de la partie. La prise de parole de Charles Kantor, le patron de cette société, était très attendue alors que l'entreprise a vécu un été chaotique. Et pour cause, trois fondateurs sur cinq ont décidé de claquer la porte de l'entreprise, seulement trois mois après sa toute première de levée de fonds de 220 millions de dollars. De quoi jeter le trouble sur ce projet ambitieux, amené à jouer dans la même cour que des acteurs comme Mistral AI.

Dans ce contexte, Charles Kantor a présenté sa vision pour tirer son épingle du jeu. Estimant que le marché de l'IA générative se répartit entre les modèles de langage, les modèles de code et les modèles d'action, l'entrepreneur a indiqué vouloir se positionner sur cette dernière catégorie mentionnée. Il estime en effet que les modèles d'action sont prometteurs puisqu'ils permettent de combiner des grands modèles de langage (LLM) avec des grands modèles de vision (VLM). Une fois que ce domaine sera parfaitement maîtrisé, H envisagera alors de s'engager dans les deux autres environnements.

Avec son approche holistique de l'IA (d'où son nom H), la startup avait indiqué à son lancement vouloir faire un pas décisif vers l'AGI (intelligence artificielle générale), qui serait capable de prendre des décisions de manière totalement autonome en prenant en compte son environnement. Une perspective très intéressante, mais qui est encore loin d'être concrète. Surtout que H n'a toujours dévoilé aucune solution. Charles Kantor s'est juste contenté d'afficher un QR code sur l'écran géant, dans l'optique d'offrir un accès aux développeurs dans quelques semaines. Il assure également qu'il y aura des annonces «*très bientôt*». Affaire à suivre !



Pas d'annonce spectaculaire mais...

A l'issue de cette deuxième édition de la conférence ai-PULSE, marquée également par des interventions de Michal Dell (en visioconférence comme Jensen Huang, le patron de Nvidia, il y a un an), Renée James

(Ampere Computing), Renen Hallak (Vast Data) ou encore Jean-Baptiste Kempf, figure française importante de l'open source et ex-CTO de Shadow (ex-Blade). L'arrivée de ce dernier chez Scaleway en tant que CTO a d'ailleurs été officialisée ce jeudi matin à Station F. Dans le cadre de ses fonctions, il sera chargé de superviser la création d'une nouvelle génération de produits et services IA souverains et européens. A noter que c'est Clara Chappaz, la secrétaire d'État en charge de l'intelligence artificielle et du numérique, qui a conclu l'événement, pour rappeler la place de choix de la France dans l'écosystème IA à l'échelle internationale alors que Paris se prépare à accueillir l'AI Action Summit les 10 et 11 février 2025.

Cette deuxième édition d'ai-PULSE contenait moins d'annonces fracassantes que la première, mais elle n'était pas moins intéressante. C'était aussi l'occasion de remettre la France sur le devant de la scène mondiale de l'IA alors qu'OpenAI, l'entreprise à l'origine de ChatGPT, s'apprête à ouvrir un bureau à Paris, dont l'inauguration est prévue le jeudi 14 novembre.

Ce cru 2024 était particulièrement technique, avec les interventions de différents acteurs de l'écosystème IA pour leur permettre d'approfondir leur vision pour le secteur, que ce soit sur les grands modèles, les GPU pour les faire tourner, l'importance d'une approche open source ou encore la réglementation. C'est le signe qu'il est peut-être enfin temps de marquer le pas après deux ans d'hystérie depuis l'arrivée de ChatGPT. On ne s'en plaindra pas... Un sentiment à confirmer l'an prochain à l'occasion de la troisième édition d'ai-PULSE.



À lire aussi
Xavier Niel réunit un casting 5 étoiles pour défendre une IA open source à la française



MADDYNEWS

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

JE M'INSCRIS

