

REGROUPEMENT 3 – APPEL À PARTICIPATION

Le Regroupement 3, dont l'objectif général¹ est d'améliorer l'adaptabilité, la fiabilité et la robustesse des grands modèles de langages (GMLs), **ouvre un appel à participation pour financer au maximum quatre projets, à hauteur de 25 000 \$ par projet pour un an.**

Les projets éligibles doivent s'inscrire dans l'un des cinq thèmes du Regroupement 3.

THÈMES

Thème 1 : Adaptation et robustesse des GMLs. Bien que les grands modèles de langage (LLM) soient des agents généralistes capables de résoudre plusieurs tâches, ils sont souvent déployés dans de nouveaux domaines pour de nouvelles tâches. L'adaptation d'un LLM à ce nouveau domaine devient rapidement importante. Cependant, il n'est pas toujours possible d'effectuer un affinage dans le nouveau domaine. Ce thème étudie l'adaptation de domaine augmentée par recherche d'information et l'adaptation culturelle des GMLs et des VLMs.

Thème 2 : GMLs modulaires. Les LLMs et les modèles vision-langage (VLM) peuvent être considérés comme des experts généralistes ou des experts orientés vers les tâches, spécialisés dans certaines modalités (langage, images, vidéos). Cependant, ils sont souvent présentés comme des modèles monolithiques, difficiles à adapter ou à intégrer, ou qui ne permettent pas une décomposition aisée des processus de raisonnement en tâches, sous-tâches, modules ou composants de connaissance. L'un des défis actuels consiste donc à créer des modèles LLM modulaires capables d'exploiter des agents, des APIs, des modules, des points de données et d'autres modèles spécifiques à une tâche et de créer des architectures réutilisables et interprétables. Ce thème se concentre sur la composition, l'inférence et les méthodes d'apprentissage servant à exploiter la modularité dans les GMLs.

Thème 3 : GMLs et VLMs non-paramétriques. Une composante importante du succès des LLM est la connaissance encodée dans leurs poids. Cependant, cette connaissance paramétrique contribue également aux défis d'adaptation des LLM et à leur non-modularité, comme souligné dans les deux points précédents, car la connaissance paramétrique conduit

¹ Pour les objectifs spécifiques et plus de détails sur le regroupement, consultez le site web d'IVADO.
<https://ivado.ca/regroupements/traitement-automatique-du-langage-naturel-taln/>

à des hallucinations, des connaissances inexactes et obsolètes, des résultats difficilement interprétables et la nécessité de réentraîner périodiquement les modèles sur de nouveaux domaines et de nouvelles tâches. Ce thème explore comment les mémoires non paramétriques basées sur des bases de connaissances peuvent être conçues et exploitées pour fournir des informations actualisées, des capacités de recherche, d'édition et de mise en cache des informations, des capacités d'agrégation des informations pour les tâches qui nécessitent une combinaison de données provenant de différentes sources et des mécanismes d'interprétabilité pour les GMLs.

Thème 4 : Interprétabilité et sécurité des GMLs. L'une des principales raisons de créer des architectures multi-agents modulaires avec des mémoires externes est de pouvoir améliorer l'interprétabilité des résultats et la sécurité des LLMs et de fournir une interface homme-machine qui augmente la confiance du grand public envers les modèles basés sur l'IA. Ce thème vise à améliorer l'interprétabilité et la fiabilité des modèles d'apprentissage automatique et à concevoir de nouvelles méthodes d'interprétabilité pour les systèmes d'apprentissage automatique.

Thème 5 : Langues autochtones et à faible ressources. Les technologies LLMs ont été limitées à l'anglais et à plusieurs langues bien dotées en ressources. Pour les langues indigènes, menacées et à faibles ressources (IEL), les LLMs n'existent pas ou sont moins performants. En outre, les variétés de langues non standards ou non prestigieuses sont souvent ignorées et négligées. Ce thème explore les techniques et méthodologies à faibles ressources pour le traitement des variétés de langues et développe des bases de données génératives collaboratives pour les langues peu étudiées afin de promouvoir la diversité linguistique et de contrer la disparition de variétés de langues.

FINANCEMENT

Le montant maximum alloué par projet est de 25 000 \$ pour un an. Ce montant est non renouvelable.

CRITÈRES D'ADMISSIBILITÉ

- La chercheuse principale ou le chercheur principal ne doit pas appartenir à l'un des regroupements de recherche IVADO
- La chercheuse principale ou le chercheur principal doit être professeure/professeur dans une université au Québec
- Le financement favorisera des projets en lien avec un ou plusieurs thèmes du Regroupement 3
- Les projets doivent être différents des projets de recherche existants dans le cadre du Regroupement 3

DATES IMPORTANTES

- **Date limite de soumission** : 15 août 2025
- **Notification des résultats** : 20 septembre 2025
- **Début prévu des projets** : 1 octobre 2025

SOUSSION

Les demandes doivent contenir les éléments suivants :

- Description du projet de recherche (max 2 pages) : contexte, objectifs, équipe, justification du budget, retombées anticipées, échéancier, considérations EDI
- CV des chercheurs et chercheuses
- CV des étudiants, étudiantes et postdocs

Les candidatures doivent être envoyées par courriel à [Danielle Maia de Souza](mailto:danielle.maia.de.souza@ivado.ca), avant le 15 août 2025, 23h59.

CONTACT

Danielle Maia de Souza, conseillère à la recherche pour le Regroupement 3 :
danielle.maia.de.souza@ivado.ca