

ARGUMENTAIRE

Demande d'autorisation préalable de recrutement

Date : 27 septembre 2023

Nom du demandeur et fonction : Yannick Estève, Professeur, Directeur du LIA

Argumentaire :

Le LIA travaille depuis sa création en 1987 dans le domaine du traitement automatique de la parole. Dans ce domaine, au cœur de l'intelligence artificielle, plusieurs paradigmes se sont succédé en 36 ans : l'ère des systèmes experts, puis celle de l'apprentissage automatique, principalement via la modélisation statistique, et aujourd'hui l'ère de l'apprentissage neuronal profond (*deep learning*).

L'opération de recherche dans laquelle s'inscrivent les missions confiées à l'agent s'intitule « Réseaux de Neurones profonds pour le Traitement de la Parole ». Elle consiste à produire des contributions scientifiques et des solutions technologiques nouvelles et transversales pour l'ensemble des travaux en traitement de la langue et de la parole du LIA.

Les compétences et expertises rares générées par les travaux de recherche du LIA dans le cadre de cette thématique de recherche sont très appréciées et demandées, avec pour conséquences de très nombreux financements en provenance d'institutions publiques et des sollicitations de la part de partenaires industriels.

Ces financements consolident les activités du LIA dans ce domaine en permettant le recrutement de doctorants et de chercheurs contractuels sous statut de post-doctorants ou ingénieurs de recherche. Cette consolidation accentue l'attractivité du LIA aux yeux des financeurs.

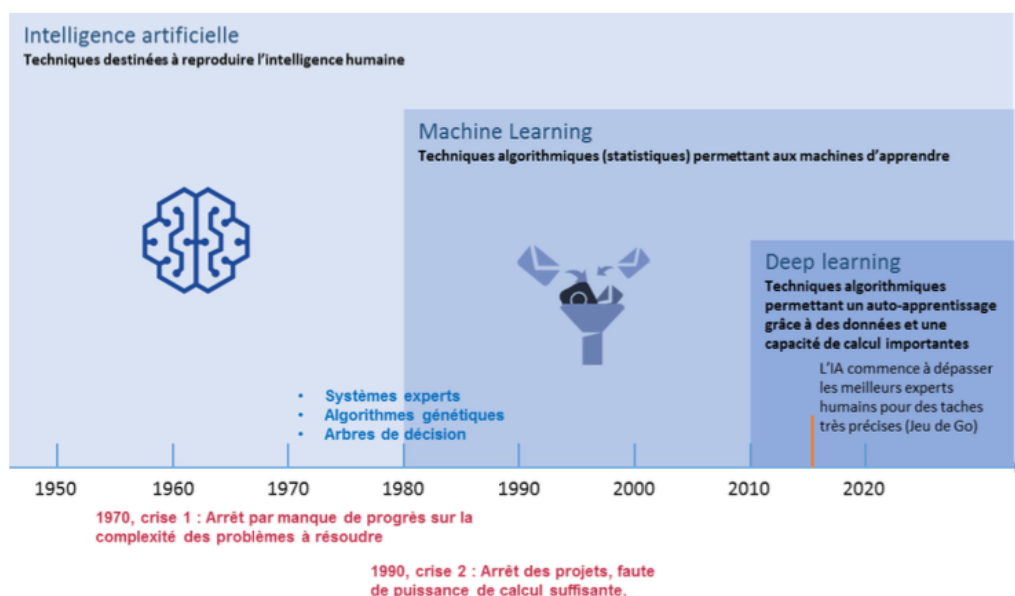
Cependant, ce modèle montre ses limites : lorsque les doctorants et post-doctorants quittent le LIA, une partie des compétences acquises dans leur travaux est perdue pour le laboratoire. En ce qui concerne les post-doctorants, leur recrutement n'apporte une réelle plus-value pour le laboratoire que lorsque la durée de leur contrat atteint deux années, la première année étant nécessaire pour une mise à niveau de la personne recrutée dans un domaine scientifique dont la vitesse d'évolution de l'état de l'art est vertigineuse.

Enfin, la très forte demande dans le secteur économique de personnels aussi qualifiés rend très difficile l'attractivité du laboratoire, en termes de salaires, mais surtout en termes de durée du contrat de travail, les entreprises privilégiant les recrutements en CDI dans ce domaine.

Pour renforcer durablement le laboratoire dans un domaine extrêmement compétitif qui ne manque pas de financements extérieurs, nous souhaiterions recruter une personne sous contrat de type CDI de mission scientifique. Cette possibilité récente nous semble particulièrement pertinente dans cette situation.

L'opération de recherche (ODR) « Réseaux de Neurones profonds pour le Traitement de la Parole » ouverte à l'automne 2019 a été abondée de plus de 300k€ en provenance de contrats industriels, ce qui illustre la capacité du LIA à récupérer des financements sur ce domaine, d'autant plus que d'autres financements encore plus élevés qui auraient pu comptablement être affectés sur cet ODR ne l'ont pas été. Il est également aujourd'hui possible d'utiliser des financements issus de projets ANR pour financer un CDI de mission scientifique. Grâce à cela, nous avons aujourd'hui une visibilité financière qui permet de garantir le financement de ce CDI de mission scientifique pour les 4 prochaines années à partir de fonds extérieurs. Nous estimons la pérennité de ce CDI de mission à environ 10 ans, et il est très probable que ces niveaux de financements seront toujours accessibles, voire accrus, dans les années à venir.

Le Pôle interministériel de Prospective et d'Anticipation des Mutations économiques (Pipame) a publié en 2019 un rapport intitulé « Intelligence artificielle - État de l'art et perspectives pour la France »¹ dans lequel a été extraite la figure suivante. Elle illustre l'évolution paradigmatique que le LIA a accompagné depuis sa création, et donne des indices sur la durée minimale à laquelle les perspectives actuelles s'accordent au sujet de l'ère du *deep learning* : l'horizon 2034 évoqué dans ce CDI de mission scientifique est un horizon tout à fait raisonnable.



Signature :

¹https://www.entreprises.gouv.fr/files/files/directions_services/etudes-et-statistiques/prospective/Intelligence_artificielle/2019-02-intelligence-artificielle-etat-de-l-art-et-perspectives.pdf

Avec l'objectif de recruter un chercheur ou une chercheuse en CDI de mission scientifique dans la thématique de l'apprentissage neuronal profond appliqué au traitement automatique de la parole, nous demandons **la prolongation de l'opération de recherche (ODR) «LES RÉSEAUX DE NEURONES PROFONDS UTILISÉS DANS LE CADRE DU TRAITEMENT AUTOMATIQUE DE LA PAROLE jusqu'au 31 décembre 2034.**

Nous demandons également le rattachement à cette ODR :

- de l'EOTP lié au contrat d'Etude C-23-LIA-02 en cours entre le LIA et la société MAIA, dont l'objet porte sur l'identification du locuteur. La recette minimale de ce contrat s'élèvera à 22.800€
- de l'EOTP à venir en lien avec la convention de collaboration dans la thèse CIFRE en cours avec la société Thalès dont le titre est « *Robustesse des modèles neuronaux pour le traitement automatique de la parole* » d'un montant de 51.743€

Enfin, toujours dans le cadre du financement du CDI de mission scientifique évoqué plus haut, **nous demandons le rattachement partiel des projets ANR suivants, sur leur volant masse salariale :**

- E-SSL « Efficient Self-Supervised Learning for Inclusive and Innovative Speech Technologies », d'un montant total de 190 685€, dont 50k€ pourront être fléchés vers cet ODR. Ce projet s'achèvera le 31 décembre 2026
- Pantagruel « Construction et évaluation de giga modèles de langue multimodaux et inclusifs (écrit, oral, pictogrammes) pour le français général et clinique », d'un montant de 140 290 € dont 129 000 € peuvent être fléchés vers cet ODR. Ce projet s'achèvera au plus tôt le 30 septembre 2026.

Avant l'ajout de ces lignes budgétaires liées à de nouveaux contrat, l'ODR « LES RÉSEAUX DE NEURONES PROFONDS UTILISÉS DANS LE CADRE DU TRAITEMENT AUTOMATIQUE DE LA PAROLE » est dotée d'un montant s'élevant à 82 324€.

Depuis 2019, les recettes de cet ODR ont été abondées par les lignes budgétaires suivantes :

EOTP NXPS1 : 30.000€
EOTP ORTP : 292.763,20€
CT SEMAXONE : 46170€
CT ATOO : 22800€
CT JURISPREDIS : 22800€
CT ZENIDOC : 10000€
CT NXPS : 58000€
CT SAFRAN : 4059,20€
CT VALIDSOFT : 19380€
CT ZENIDOC2 : 63954€
AVT 1 VALIDSOFT : 22800€
AVT 2 VALIDSOFT : 22800€

TOTAL : 322.763,20€

Yannick Estève
Directeur du LIA