



Post-doctorante ou Post-doctorant : « Pour une IA soutenable : étude du coût calcul de l'apprentissage profond » – CDD de 18 mois

Environnement du poste :

L'Institut Mines-Télécom est le 1er groupe public de Grandes Écoles d'ingénieur.e.s et de management de France. Constitué de huit Grandes Écoles publiques et de deux écoles filiales, l'Institut Mines-Télécom anime et développe un riche écosystème d'écoles partenaires, de partenaires économiques, académiques et institutionnels, acteurs de la formation, de la recherche et du développement économique.

L'Institut Mines-Télécom donne une place centrale à la souveraineté et la sobriété numérique dans sa stratégie, fort de sa communauté scientifique Data&IA œuvrant entre autres pour une IA soutenable. Cet axe stratégique a amené l'Institut Mines-Télécom à rejoindre le réseau d'excellence ENFIELD, qui réunit trente partenaires européens autour de travaux pour une IA de confiance et verte (*European Lighthouse to Manifest Trustworthy and Green AI*, voir <https://www.enfield-project.eu/>). C'est dans le cadre d'ENFIELD qu'est ouvert un poste de post-doc à Mines Saint-Étienne, l'une des huit écoles de l'Institut Mines-Télécom.

Au sein de Mines Saint-Étienne, c'est l'Institut Fayol et le département Informatique et systèmes intelligents (ISI), rattaché au Laboratoire d'informatique, de modélisation et d'optimisation des systèmes (LIMOS, UMR 6158), qui portent la thématique de l'IA soutenable, sous l'angle du coût calcul de l'apprentissage et de l'analyse d'impact environnemental d'un système d'IA.

Enjeux scientifiques :

Les modèles de langage et autres modèles fondationnels basés sur des réseaux de neurones profonds sont devenus incontournables en quelques années. Cependant, dès l'introduction de l'architecture Transformer dans la littérature s'est posée la question du coût calcul de ces modèles. Leur grande capacité à généraliser tient principalement au fait qu'ils possèdent un très grand nombre de paramètres configurables, comparés à leurs prédécesseurs. Or, on peut observer empiriquement que les progrès effectués depuis 2012 en apprentissage statistique s'accompagnent d'une croissance exponentielle du coût calcul nécessaire au développement d'un modèle.

L'adoption croissante des réseaux de neurones dans les systèmes d'information pose le problème de la maîtrise du coût calcul de l'IA et son incidence sur l'impact environnemental du numérique. Maîtriser le coût calcul de l'IA appelle en retour à développer des indicateurs solides pour prendre des décisions en amont de la conception d'un système d'information. L'éco-conception logicielle impose par exemple de justifier qu'une solution coûteuse remplit un besoin précis et n'est pas substituable.

Or, actuellement, il existe peu d'indicateurs *a priori* du coût calcul d'un modèle d'apprentissage. De nombreux paramètres entrent en jeu, comme la quantité de données d'entraînement, la complexité du modèle ou les caractéristiques matérielles de la plateforme sur laquelle le modèle est évalué. Certains de ces paramètres sont disponibles en ligne : Hugging Face publie par exemple des « *model cards* » indiquant combien le modèle a de paramètres, quelles données ont été utilisées pour l'apprentissage et, parfois, une estimation de la quantité d'énergie consommée. En revanche, une *model card* ne fournit aucune donnée sur le modèle après entraînement (par exemple, pour faire des inférences).



Les pratiques actuelles autour de l'IA appellent à développer un indicateur de coût calcul répondant aux exigences de l'éco-conception. Cet indicateur pourra être évalué sur différents modèles développés par les partenaires d'ENFIELD, ainsi que par la communauté scientifique Data&IA de l'Institut Mines-Télécom.

Missions principales :

La mission principale du ou de la post-doc sera de concevoir une méthode d'analyse statique (avant exécution) pour estimer le coût calcul d'un modèle d'apprentissage implémenté à partir d'une bibliothèque de calcul vectoriel comme PyTorch, TensorFlow ou NumPy. La méthode devra prendre en compte des paramètres de taille de données, de taille de modèle mais aussi de caractéristiques matérielles comme la taille de mémoire disponible et le nombre de processeurs disponibles en parallèle.

Le ou la post-doc aura par ailleurs pour mission de co-animer les échanges scientifiques entre partenaires d'ENFIELD, qui se structurent de façon matricielle en piliers (*Green AI*, *Adaptive AI*, *Human-Centric AI* et *Trustworthy AI*) et en domaines d'application (*Energy*, *Healthcare*, *Manufacturing* et *Space*). Chaque pilier et chaque domaine d'application est associé à un groupe de travail susceptible d'échanger avec les autres.

Activités principales :

Le post-doc comprend les activités suivantes :

- Revue des modèles disponibles sur Hugging Face, des modèles développées à l'IMT et des modèles développés par les partenaires d'ENFIELD,
- Conception d'un indicateur de coût calcul pour ces modèles (par exemple en nombre de FLOP),
- Développement d'un outil d'analyse statique pour fournir des estimations de coût calcul à partir de code (par exemple, à partir des fonctions de la bibliothèque PyTorch),
- Participation aux groupes de travail d'ENFIELD,
- Participation aux actions de la communauté scientifique Data&IA de l'Institut Mines-Télécom.

Profil :

Connaissances, savoirs :

- Diplôme de niveau Bac+8 en intelligence artificielle, mathématiques appliquées ou génie logiciel

POURQUOI NOUS REJOINDRE :

Le réseau d'excellence ENFIELD permet :

- L'acquisition d'expérience sur un grand projet européen (Horizon Europe),
- Des rencontres avec les membres de trente instituts de recherche et universités européennes (Norvège, Portugal, Espagne, Italie, Allemagne, Autriche, Danemark, etc.),
- Un contact avec des chercheur-es dans de nombreux sous-domaines de l'IA.



L'Institut Mines-Telecom se caractérise par :

<https://www.youtube.com/watch?v=m39m6hdNC48>

- Un environnement scientifique d'excellence,
- Un groupe présentant des entités sur l'ensemble du territoire français.

Mines Saint-Etienne se distingue par :

- Un environnement de travail privilégié avec un taux d'encadrement des étudiants et un taux d'environnement (fonctions support et soutien) élevé,
- Des moyens expérimentaux et numériques de premier plan,
- Une activité de recherche contractuelle importante (11 M€ / an de contrats en Recherche et Innovation), majoritairement avec des partenaires industriels,
- 25% d'étudiants internationaux, Membre du réseau T.I.M.E. et de l'Université Européenne EULIST,
- Un centre de Culture Scientifique Technique et Industrielle – La Rotonde - unique en France, qui démultiplie son impact sur la société (> 50 000 visiteurs par an),
- Cadre de vie agréable : proximité du parc naturel du Pilat, faible coût de la vie, troisième métropole en région Auvergne-Rhône-Alpes et ville créative UNESCO
- Frais de transports en commun pris en charge à hauteur de 75% (sous conditions),
- Forfait mobilité durable,
- Foyer du personnel (activités sportives, culturelles, avantages CE sur des loisirs et temps de convivialité),
- Possibilité de télétravail partiel,
- 49 jours de congés et RTT.

Conditions de recrutement :

- CDD de droit public de 18 mois
- Prise de fonction souhaitée : **entre juin et septembre 2024**
- La rémunération sera fixée selon le profil du candidat, en fonction des règles définies par le cadre de gestion de l'Institut Mines Télécom
- Temps plein
- Poste basé à Saint-Étienne

Le poste est ouvert à toutes et tous avec, sur demande, des aménagements pour les candidates et candidats en situation de handicap.

Modalités de candidature :

Les dossiers de candidature (CV, lettre de motivation, lettre de recommandation le cas échéant) sont à déposer sur la plateforme RECRUITEE **au plus tard le 08 mai 2024** :

<https://institutminestelem.com/o/post-doctorante-ou-post-doctorant-pour-une-ia-soutenable-etude-du-cout-calcul-de-lapprentissage-profond-cdd-de-18-mois-2>

Dans le cadre de sa politique Égalité, Diversité et Inclusion, l'École des Mines de Saint Etienne est un employeur soucieux de l'équité de traitement entre les candidatures.

Pour en savoir plus :

Pour tous renseignements sur le poste, s'adresser à :



Victor Charpenay (maître de conférences), victor.charpenay@emse.fr

Olivier Boissier (directeur de l'institut Fayol), olivier.boissier@emse.fr

Pour tout renseignement administratif, s'adresser à :

Julie Jaffre, julie.jaffre@emse.fr, 04 77 42 00 17

Information des candidats sur le traitement des données personnelles :

<https://www.mines-stetienne.fr/wp-content/uploads/2018/12/Informations-des-candidats-sur-les-traitements-de-donn%C3%A9es-personnelles.pdf>