

Université de technologie de Compiègne - Proposition de thèse

1^{re} partie : Fiche scientifique	
Intitulé de la thèse	Du paradigme de l'inférence au paradigme de l'apprentissage : enjeux épistémologiques du <i>deep learning</i> dans la recherche en biologie et en biomédecine
Type de financement	Contrat doctoral sur allocation Ministère
Laboratoire d'accueil	unité de recherche : COSTECH équipe de recherche : CRED site web :
Directeur(s) de thèse	Xavier GUCHET (Costech, HDR) et Pierre STEINER (Costech, HDR)
Domaines de compétence	Sciences de la Société Sciences humaines
Description du sujet de thèse	<p>Les sciences cognitives et l'informatique ont connu un développement conjoint. Les connaissances sur le cerveau et la cognition humaine ont servi de modèles pour la construction de machines informatiques, et inversement, les nouvelles technologies ont servi de modèles pour penser la cognition humaine. L'inférence (sous forme déductive) et l'apprentissage (par induction) sont les deux grands paradigmes qui ont façonné les sciences cognitives et l'informatique, impliquant deux orientations très différentes pour le développement des intelligences artificielles (IA).</p> <p>L'IA classique s'est construite à partir du postulat que ce qu'il s'agit de transposer dans la machine, c'est le caractère calculatoire de la pensée. Toutefois, le paradigme de l'inférence (la pensée comme calcul) ne permet pas de saisir l'ensemble des dimensions de l'intelligence humaine, notamment en ce qui concerne sa capacité d'adaptation, d'improvisation, de prise d'initiative et de flexibilité. L'apprentissage résulte soit d'une programmation par une instance externe au système, soit de la déduction de résultats qui intègrent une base de connaissances.</p> <p>Les découvertes sur le système nerveux ont fait émerger un nouveau paradigme : l'apprentissage par induction. L'apprentissage est en effet ici défini comme la modification d'un comportement à la suite d'expériences répétées. Le paradigme de l'apprentissage est à l'origine d'une nouvelle lignée technique : les IA apprenantes, notamment le <i>machine learning</i> et le <i>deep learning</i> lorsque ces algorithmes utilisent des réseaux de neurones. On parle d'apprentissage dans la mesure où la machine sélectionne, essaye après essai, les mouvements ou les actions qui la rapprochent de son objectif, et construit une marche à suivre afin d'y parvenir. Au lieu d'exécuter des séries d'instructions, les IA apprenantes sont capables de se paramétrer elles-mêmes en fonction des données qui leur sont fournies.</p> <p>La machine conçue selon le paradigme de l'inférence n'avait qu'un rôle facilitant, elle permettait un gain de temps ou un gain de simplicité, mais ne pouvait pas acquérir par elle-même de nouvelles connaissances ou capacités générales. L'IA apprenante doit quant à elle découvrir par elle-même des régularités et des connexions stables dans des jeux de données, là où les outils classiques ne le permettent pas ou bien très difficilement. Le programme de <i>deep learning</i> AlphaFold, utilisé pour prédire la structure 3D des protéines à partir de leur séquence d'acides aminés (problème dit du repliement des protéines), en est une bonne illustration.</p> <p>Le <i>deep learning</i> est amené à jouer un rôle majeur dans la recherche scientifique, précisément parce qu'il permet de résoudre des problèmes inductifs (prédire, généraliser, dégager des invariants et des structures) dont la solution avec les technologies classiques ne peut se faire qu'en mobilisant une puissance de calcul dont personne ne dispose, ainsi qu'un temps très</p>

	<p>long.</p> <p>Ce projet de thèse, croisant la philosophie des techniques, les sciences cognitives et l'histoire des sciences et des techniques du vivant, entend étudier l'implantation du <i>deep learning</i> dans les sciences, en se focalisant plus précisément sur la biologie et la recherche biomédicale. La littérature sur les utilisations du <i>deep learning</i> dans la recherche scientifique est déjà importante, mais elle traite essentiellement des promesses et des bénéfices escomptés du <i>deep learning</i>. En revanche, il n'existe que très peu de travaux portant sur les usages effectifs du <i>deep learning</i>, notamment dans des laboratoires de biologie et de biomédecine. Les études empiriques font défaut. Il s'agira dans la thèse d'examiner la façon dont l'utilisation du <i>deep learning</i> reconfigure le cadre épistémologique des sciences du vivant, notamment leurs modes de production et de validation des connaissances.</p> <p>La thèse sera amenée à répondre aux trois questions suivantes :</p> <p>-1 L'utilisation de l'informatique en biologie prend son essor dans les années 1960 (Strasser 2011). En quoi le <i>deep learning</i> introduit-il, ou pas, une rupture épistémologique dans l'histoire longue des outils d'aide à la résolution de problèmes en biologie ? Cette mise en perspective historique devra permettre de cerner avec précision ce qu'il y a de spécifique, épistémologiquement parlant, dans ces nouveaux dispositifs.</p> <p>-2 Comment le problème du défaut « d'explicabilité » des algorithmes de <i>deep learning</i> impacte-t-il la recherche en biologie et en biomédecine ? Les programmes de <i>deep learning</i>, à base de réseaux de neurones, sont en effet très complexes et il n'est pas toujours possible de retracer le cheminement qu'ils ont suivi. Comment les biologistes s'arrangent-ils avec ce problème ?</p> <p>-3 Comment caractériser ces dispositifs de <i>deep learning</i> au regard de l'utilisation qu'en font les biologistes ? Peut-on encore parler « d'outils », selon une perspective de <u>suppléance</u> du biologiste, ou bien faut-il considérer que ces dispositifs ont vocation à remplacer les biologistes dans l'exécution de certaines tâches, selon une logique de <u>substitution</u> ?</p> <p>La thèse aura une forte dimension empirique. Elle s'appuiera sur des observations in situ et sur des entretiens semi-directifs avec des biologistes et des ingénieurs concepteurs des dispositifs de <i>deep learning</i>, en vue de mieux connaître les usages effectifs de ces dispositifs dans les pratiques scientifiques.</p>
Mots clés	Biologie, Biomédecine, <i>Deep learning</i> , Épistémologie, Intelligence artificielle, Philosophie
Profil et compétences du candidat	Spécialisation en philosophie des sciences, philosophie des sciences cognitives et/ou philosophie des techniques, attestée par un ou plusieurs mémoires de Master (1 et 2) dans ces domaines.
Date de début de la thèse	Le plus tôt possible
Lieu de travail de thèse	UTC

2 ^e partie : Fiche de poste	
Durée	36 mois
Possibilité missions complémentaires	Enseignements TSH (Épistémologie, Philosophie des sciences, Philosophie des techniques)
Laboratoire d'accueil	COSTECH. Expertises en Épistémologie et Philosophie des sciences, Philosophie des techniques, Sciences cognitives, Philosophie de l'esprit
Moyens matériels	Bureau collectif, un ordinateur
Moyens humains	35 EC (dont deux PAST), 24 doctorants, 5 post-doctorants, 2 BIATSS/ITA
Moyens financiers	Pour les terrains biomédecine, le projet MaLO – Machine Learning in Oncology (financé par l'Institut National du Cancer) pourra financer les missions
Modalités de travail	Implication dans l'axe Care Technologies de Costech ; implication dans le projet MaLO (organisation de séminaires) ; réunions mensuelles avec les dir. de thèse
Projet de recherche lié à cette thèse	Projet MaLO
Collaboration(s) nationale(s)	Partenaires du projet MaLO (Université Lyon III ; Université de Nantes ; Université Catholique de Lille ; CNRS Délégation Bretagne)
Collaboration(s) internationale(s)	Université Catholique Louvain la Neuve
Thèse en cotutelle internationale	Non
Coordonnées de la personne à contacter	xavier.guchet@utc.fr ; pierre.steiner@utc.fr

Contactez d'abord le directeur de thèse avant de renseigner un dossier de candidature en ligne sur <https://webapplis.utc.fr/admissions/doctorants/accueil.jsf>