

Appel à manifestation d'intérêt - Chaire de professeur junior
Fiche projet type

Établissement/organisme porteur : Inria

Nom du chef d'établissement/d'organisme : Bruno Sportisse

Site concerné : Centre Inria Nancy – Grand Est

Région académique : Grand Est

Établissements/organismes partenaires envisagés : le cas échéant

Nom du projet : IA pour la simulation d'interactions moléculaires en vue de la conception d'inhibiteurs à portée thérapeutique

Acronyme : IA-SIMIT

Mots-clés : Simulation numérique, interactions moléculaires, intelligence artificielle, molécules thérapeutiques, inhibiteurs.

Durée visée : 5 ans

Thématique scientifique : Biologie Numérique

Section (s) CNU/CoNRS/CSS correspondante (s) : Sections CNU 27, 64, CoNRS 6, 20, 13, 16

Stratégie d'établissement :

Cette chaire vise à soutenir et intensifier les travaux de CAPSID, l'une des 12 équipes-projets Inria dans le domaine de la Biologie Numérique, en cohérence avec la politique de site du centre de Nancy. Ce faisant, elle contribue à deux objectifs centraux de la stratégie scientifique d'Inria : i) développer et intensifier les recherches à la frontière entre le numérique et les autres disciplines, afin de permettre des avancées scientifiques majeures et l'émergence à terme de nouveaux domaines scientifiques encore inexplorés, et ii) accroître la maturité des développements technologiques associés au-delà de la preuve de concept grâce à une connexion étroite avec les utilisateurs de ces technologies, afin d'avoir un impact concret sur les deux disciplines et de fournir une base solide pour des transferts industriels ou des usages cliniques. En parallèle, Inria mise sur ces nouveaux liens scientifiques pour construire et proposer des formations originales alliant des connaissances fondamentales solides et la pratique expérimentale dans plusieurs disciplines. La chaire contribue aussi à la stratégie de l'établissement en matière d'attractivité internationale.

Stratégie du laboratoire d'accueil :

La chaire vise à renforcer l'équipe-projet CAPSID, commune avec le CNRS et l'Université de Lorraine au sein du Laboratoire Lorrain de Recherche en Informatique et ses Applications (LORIA, UMR 7503), et à intensifier sa collaboration avec le Laboratoire de Physique et Chimie Théoriques (LPCT, UMR 7019) sur l'intégration de résultats expérimentaux dans les simulations numériques en biophysique d'une part, et avec le laboratoire d'Ingénierie Moléculaire et Physiopathologie Articulaire (IMoPA, UMR 7365) sur la validation de molécules actives et leurs applications thérapeutiques d'autre part. Cet objectif s'inscrit dans la volonté du Centre Inria Nancy – Grand Est de contribuer pleinement à l'I-Site pérennisé Lor-

rainé Université d'Excellence construit autour de la thématique de l'« ingénierie systémique ». La stratégie du centre consiste à susciter, accompagner et développer la pluridisciplinarité autour du numérique, en soutien des axes forts du site, parmi lesquels le domaine de la santé. D'autre part, le volet formation de cette proposition viendra renforcer le dispositif ORION « Oser la Recherche durant la formation ».

Résumé du projet scientifique :

Le projet scientifique associé à la chaire se situe à l'intersection des domaines de la bioinformatique structurale, de la simulation numérique de systèmes biophysiques et de l'intelligence artificielle (IA), pour des applications au service de la santé. Il s'agira de proposer, mettre en œuvre et évaluer des approches innovantes, à base d'IA, pour la conception de molécules thérapeutiques nouvelles, capables de mimer et/ou inhiber des interactions protéiques, telles que les interactions virus-récepteur qui permettent la propagation des virus. La conception de méthodes d'IA, orientées par les données (data-driven) et adaptées pour la modélisation et le design d'interactions moléculaires, repose en particulier sur les méthodes d'apprentissage automatique par réseau de neurones profonds, sur l'apprentissage de représentation dans un espace latent par les auto-encodeurs, et sur la modélisation des interactions par des graphes moléculaires. Les défis à relever sont nombreux: depuis l'accélération des programmes de simulation dynamique des interactions moléculaires, la modélisation de la flexibilité des molécules biologiques (protéines, peptides) par des méthodes d'ensemble, les calculs d'énergie libre de liaison à partir des simulations, jusqu'à la prédiction de molécules actives par leur propriété d'inhiber des interactions nocives.

Résumé du projet d'enseignement :

Le projet d'enseignement associé à la chaire aura une coloration pluridisciplinaire soutenue. Il s'agira de toucher à la fois les étudiants des filières informatiques et biologiques, en apportant aux premiers des applications empruntées à la biophysique ou à la bioinformatique structurale, avec des enjeux en santé, et aux seconds des notions d'informatique essentielles pour comprendre les méthodes de l'IA utilisées dans divers domaines d'étude et de conception de molécules thérapeutiques. Les filières concernées déjà sensibilisées à ce type de pluridisciplinarité à l'Université de Lorraine pourront être le parcours IAMD (« Intelligence Artificielle et Masses de Données ») des élèves ingénieurs de Telecom-Nancy, le module « Biologie Moléculaire et Applications » aux Mines, département Génie Industriel et Mathématiques Appliquées, le Cursus Master-Ingénieur Biologie-Santé-Environnement (CMI BSE), l'option Option Ingénierie Moléculaire du Master de Biotechnologies, etc. L'objectif à moyen terme est de faire émerger un enseignement commun à deux filières ou plus, afin de favoriser le dialogue entre les disciplines.

Synthèse financière :

En sus de la rémunération du titulaire du la CPJ, les besoins financiers nécessaires pour mener à bien le projet scientifique incluent :

- 1 doctorant sur 36 mois (120 k€)
- 1 post-doctorant sur 24 mois (96 k€)
- des frais de fonctionnement : ordinateur portable, gratifications de stages, missions (34 k€)

soit un coût total de 250 k€. Un co-financement de 50 k€ sera demandé à la Région Grand Est dans le cadre du dispositif existant de co-financement de thèses en IA. Nous sollicitons un financement de 200 k€ de l'ANR.

Total financé sur CPJ (dont package ANR)	498 k€
Co-financement	87 k€
Total du projet	585 k€

Diffusion scientifique : Les résultats seront diffusés dans des revues de premier plan dans le domaine de la Biologie Numérique, par exemple :

- Nature Scientific Reports
- PLOS Computational Biology
- Proteins
- Journal of Chemical Information and Modeling

Le titulaire de la CPJ participera au challenge CAPRI (*Critical Assessment of Prediction of Interactions*) : <https://www.ebi.ac.uk/pdbe/complex-pred/capri/>.

Science ouverte : Le projet s'inscrit dans une démarche de science ouverte, qui pourra être mise en œuvre selon les principes suivants.

- Mise en œuvre d'un PGD (Plan de Gestion des Données) sur le modèle Opidor
- Mise à disposition des jeux de données dans DOREL (Données de la Recherche à l'Université de Lorraine) avec les métadonnées nécessaires, pour l'obtention d'un identifiant pérenne de type DOI.
- Mise à disposition des codes et des logiciels sur une plateforme publique de type GIT avec licence logiciel libre associée (type MIT). Pour les logiciels aboutis, enregistrement dans l'initiative Software Heritage.
- Publications dans des journaux Open Access et enregistrement dans HAL avec mise à disposition d'une copie pdf auteur.

Science et société : Le titulaire de la CPJ et les autres personnes impliquées interviendront dans des manifestations grand-public telles que la Fête de la Science, le Salon Cité-Santé, ou le colloque IA et Santé organisé chaque année par le CHRU de Nancy, pour expliquer le projet et faire connaître les approches innovantes en cours de développement. A l'issue du projet, ce seront des résultats qui pourront être communiqués.

Indicateurs : Les indicateurs de suivi pourront être :

- Présentation du PGD
- Recrutement d'un doctorant et d'un post-doctorant
- Encadrement de qualité, attesté par les publications/communications du doctorant et du post-doctorant
- Mise à disposition effective des logiciels et des données
- Collaboration effective entre le titulaire de la chaire et les laboratoires LPCT et IMoPA, attestée par des co-encadrements ou des publications communes
- Soutenance HDR du titulaire de la CPJ
- Présentation de la maquette d'enseignement commun à deux filières ou plus.

Un espace partagé de collaboration sera ouvert sur mybox.inria.fr ou sur EDC à l'Université de Lorraine afin d'y déposer tous les documents relatifs au projet et de pouvoir suivre les indicateurs. Un comité de suivi de projet sera organisé tous les ans pour suivre l'avancement du projet, avec tous les participants au projet et des représentants des tutelles concernées.