Chaire de professeur junior - Fiche projet type

Etablissement/organisme porteur : Inria
Nom du chef d'établissement/d'organisme : Bruno Sportisse
Site concerné : *Lyon*
Région académique : Auvergne-Rhône-Alpes

Établissements/organismes partenaires : Centre Léon Bérard/Université Claude Bernard Lyon 1

Nom du projet :
Cancer: Artificial Intelligence for Multiomics ANalyses
Acronyme : **CAIMAn**

Mots-clés : *donner 5 mots-clés caractérisant le projet scientifique*
Intelligence Artificielle; Oncologie; Données MultiOmiques; Médecine de Précision; Next Generation Sequencing.

Durée visée : 5 ans

Thématique scientifique : Santé Numérique

Section (s) CNU/CoNRS/CSS correspondante (s) : CNU 26/27, CoNRS 6/41- CSS 5/6

Stratégie d'établissement : *décrire en quoi le recrutement est en lien avec la stratégie de l'établissement (15 lignes maximum)*

En dépit des progrès dans leur prévention, leur compréhension et leur prise en charge, les cancers restent la cause majeure de mortalité en France avec 382 000 nouveaux cas chaque année et 150 000 décès. On estime que plus de 60% des Français présenteront un cancer au cours de leur vie. La compréhension croissante de la base génomique et immunologique du cancer menant au développement de la médecine de précision a permis une amélioration notable de la prise en charge de plusieurs types de cancer et le nouveau Plan Cancer français à 10 ans vise à amplifier ce mouvement, suscitant également une nouvelle impulsion dans le champ de la prévention dans tous ses aspects.
La Santé Numérique constitue un domaine d'impact sociétal majeur dans lequel Inria s'investit activement au niveau national depuis plusieurs années, au travers notamment du programme de création d'équipes communes Inria-Inserm.
Le site Inria de Lyon, créé en 2021, a placé cet axe dans ses priorités les plus élevées et vise, à terme, à la création de nouvelles équipes communes au sein de l'écosystème médical lyonnais. Cet écosystème se décline actuellement autour de trois grands axes dans lesquels Lyon est leader aux niveau national et international : l'infectiologie/virologie; l'immunologie, et la cancérologie. Si le centre Inria de Lyon a historiquement déjà noué des liens avec les deux premiers domaines, notamment à travers la création d'équipes de modélisation, les contacts avec la cancérologie restent encore peu formalisés et constituent un enjeu stratégique majeur du centre.
Les compétences développées dans le domaine de l'apprentissage, de la gestion des grands volumes de données et de l'intelligence artificielle viendront compléter la modélisation et la simulation déjà présentes au sein des équipes existantes pour développer, dans le cadre de la cancérologie, de nouvelles approches scientifiques au sein d'une équipe commune avec le centre Léon Bérard. Un support important au développement d'outil logiciel sera également au cœur du projet.

Stratégie du laboratoire d'accueil : *décrire en quoi le recrutement est en lien avec la stratégie du laboratoire d'accueil (15 lignes maximum)*

Lyon dispose d'un ensemble unique d'expertises, de programmes d'enseignement et de production scientifique en oncologie, rassemblant le plus grand ensemble de chercheurs en cancérologie en France, localisé principalement au Centre de Recherche en Cancérologie de Lyon (CRCL), des équipes expertes

de recherche clinique en milieu hospitalier universitaire, un centre de cancérologie complet **Léon Bérard** (**CLB**) et le Centre international de recherche contre le cancer de l'OMS.
Lyon est un site d'excellence pour la recherche sur le cancer, comme l'attestent à la fois :
- le LAbEx « DEVweCAN », qui vise à lier progression tumorale et réactivation des processus embryonnaires,
- l'Institut Convergence PLAsCAN et
- la fondation de coopération scientifique Synergie Lyon-Cancer.
Le CLB est également un partenaire clef du groupement de coopération sanitaire (GCS) Auragen qui constitue l'une des deux plateformes nationales sélectionnées dans le cadre du Plan France-Médecine Génomique 2025 pour la mise en œuvre de la médecine de précision (séquençage de génomes entiers) en maladies rares et en oncologie.
Le CLB, avec ses partenaires Hospices civils de Lyon (HCL) et Université Claude Bernard Lyon 1 (UCBL), dispose également d'un des deux premiers « Sites de Recherche Intégrée sur le cancer(SIRIC), le LYriCAN, est le site d'un centre de phase I labellisé enfant et adulte capable de mettre en place de études de premières administration a l'homme, et dispose de réseaux d'excellences sur le plan international sur le thème des tumeurs rares, pilotant par exemple l'*European Reference Network for rare adult cancers* (EURACAN), et d'un réseau national unique à travers la présidence du réseau Unicancer, fédération des centres de lutte contre le cancer.
L'ambition du CLB est de favoriser des programmes de recherche translationnelle innovants, de nouveaux paradigmes pour la compréhension de la biologie du cancer, basés sur l'utilisation intensive des données cliniques et omiques visant au développement d'une médecine de précision universelle.

Résumé du projet scientifique : *15 lignes maximum*

Si la plupart des applications phares de l'IA en oncologie concernent essentiellement le domaine de l'imagerie médicale (radiomique) ou le traitement du langage naturel (NLP), à visée essentiellement prédictive, nous souhaitons ici explorer une dimension différente (mais complémentaire), celle des données multi-omiques (génomiques, transcriptomiques, métabolomiques, ...) dans laquelle elle reste encore insuffisamment développée. Le CLB est déjà fortement investi dans le recueil et la structuration de ces données à travers notamment la participation aux projets Consore (extraction automatique de données cliniques) et Osiris (structuration des données cliniques et génomiques), la mise en place d'un datalab et la plateforme de bioinformatique Gilles Thomas qui a développé une expertise internationale (notamment au travers des projets de l'International Cancer Genome Consortium) dans le domaine du traitement de données de génomes complets.
Le projet porte donc sur le développement de méthodologies du domaine de l'IA (Deep Learning) visant à la mise en correspondance des informations cliniques (pathologie, réponse à un traitement, ...) et multi-omiques (altérations génomiques, expression génique, ...). L'application principale porte donc sur l'annotation de la pathogénicité des altérations génomiques ou d'expression observées. Elle vise, concrètement, à fournir des outils permettant d'assister les biologiste-génomicien et les cliniciens durant les réunions de concertation pluridisciplinaires (RCP). Il est important de noter qu'il ne vise pas à les remplacer mais à les "augmenter" c'est à dire à leur permettre de se concentrer sur les résultats les plus significatives du point de vue biologique ou clinique. De ce point de vue, le projet se place donc dans une perspective prédictive mais également cognitive. Une attention particulière devra être apportée à l'interprétabilité des modèles, condition cruciale dans ce domaine pour acquérir la confiance des praticiens. Enfin, le projet Auragen, dans le cadre de France Médecine Génomique 2025, constituera un champ applicatif direct pour ces nouveaux outils.

Du point de vue organisationnel le projet vise à la création d'une **équipe commune Inria-CLB** dédiée à ces méthodologies et s'appuyant sur le recrutement de deux permanents :
- la chaire junior qui fait l'objet de cette demande et
- un chercheur plus senior qui fera l'objet d'un AO piloté par la fondation Synergie-Lyon Cancer.

Un search committee conjoint Inria-CLB-Université sera mis en place pour ces deux recrutements.

Résumé du projet d'enseignement : *15 lignes maximum*

La personne recrutée participera aux enseignements liés à la santé de précision dans le cadre du parcours *Diagnostic de précision et médecine personnalisée* de l'UCBL en y apportant son expertise spécifique (sciences des données, intelligence artificielle). Ce programme est ouvert aux médecins, vétérinaires, pharmaciens et doctorants en sciences biologiques, mathématiques, épidémiologiques , ainsi qu'aux (bio)informaticiens afin de former la prochaine génération de praticiens disposant d'un double expertise

permettant la pénétration de ce corpus de connaissance dans la population des jeunes praticiens et docteurs en sciences en formation.
L'enseignant mettra en place des modules de formations au sein des cursus médicaux, pharmaceutiques, vétérinaires sur ces thèmes, développera un parcours de Master en Biologie et Santé (M2) et un programme doctoral, ouvert à toute la communauté scientifique nationale, aux étudiants en formation tout comme aux praticiens en exercices qui devront intégrer ces nouveaux paradigmes.
L'enseignement sera structuré autour de séminaires interactifs sur les technologies et les approches de pointe actuelles (bioinformatique, IA, modélisation mathématique, thérapies innovantes, économie et éthique de la santé, …).
Un enseignement spécifique post doctoral type Diplôme inter-universitaire (DIU) pourra être développé au sein des 18 centres de cancérologie Unicancer et des sites universitaires intéressés afin de piloter cette offre de formation stratégique pour les praticiens qui devront l'intégrer à leurs futures pratiques.

Synthèse financière : *à réaliser à partir de la fiche financière jointe, décrire les besoins financiers et leur répartition pour mener à bien le projet scientifique (doctorant, post-doctorant, IT, équipement, ...)*

Total financé sur CPJ (dont package ANR)	962 k€
Co-financement	720 k€
Total du projet	1,682 k€

Diffusion scientifique : *préciser les résultats attendus en termes de diffusion scientifique (publications, communications, ...)*

Il est attendu du/de la titulaire de la chaire de valoriser les travaux et résultats obtenus à la fois dans des revues à caractère méthodologique (en informatique) et à caractère biologique et clinique. Les cibles pour la seconde catégorie sont des journaux de rang A et des congrès internationaux de référence en Oncologie (ESMO -European Society for Medical Oncology-, ASCO -American Society of Clinical Oncology-, …).
Le caractère transdisciplinaire de la recherche menée doit conduire à la publication dans des revues généralistes à fort impact (Nature, Science).
Il sera également important de veiller à la diffusion des connaissances auprès des soignants, des patients et des associations de patients et du grand public.

Science ouverte : *le projet s'inscrit-il dans une démarche de science ouverte ? Si, oui décrire sa mise en œuvre.*

Il est attendu du projet une modification des pratiques des oncologues qui soit mesurable : dans un premier temps au niveau du Centre Léon Bérard, et rapidement au sein des 18 centres Unicancer, mais qui devra ensuite diffuser plus profondément dans cette communauté.
A cet effet, les développements méthodologiques réalisés devront être implémentés dans des outils (classificateurs) à destination des communautés cibles. Les codes sources devront être mis à disposition à travers les outils dédiés à cet usage (gitlab). Le développement d'interfaces dédiées aux professionnels de santé (portails web) sera envisagé dans un second temps en lien avec la Direction des services informatiques (DSI) du centre Léon Bérard et, éventuellement, un prestataire spécialisé. Les données produites par les recherches seront également mises à disposition dans le respect des normes en vigueur (RGPD).

Science et société : *le projet envisage-t-il une communication auprès du grand public ? Si oui : préciser de quelle manière et à quelle échéance.*

La communication en direction du grand public peut immédiatement être envisagée au niveau des : soignants et patients (via les Sites Web CLB, Unicancer, les médias de diffusion interne du CLB et l'organisation de séminaires ouverts); associations de patients (via les sites de ces associations, les réseaux sociaux ainsi que par la participation à des séminaires dédiés). Enfin, dans un second temps, un

plus large public pourra être touché au travers des organisations nationales (INCa) et réseaux et sociétés savantes internationales (EURACAN, ESMO) qui sont très sensibilisés à cette communication.

Indicateurs : *préciser les indicateurs de suivi du déploiement du projet et la méthodologie de leur suivi*

Indicateurs prioritaires	
Précoces (<2 ans)	1. Recrutement des ingénieurs, doctorants et post-doctorants constituant la chaire 2. Participation à la soumission de projets « santé » d'envergure nationale (ANR, INCa, FRM…)
Moyen terme (2-5 ans)	1. Publications dans des journaux scientifiques généralistes à fort impact 2. Mise en place d'un enseignement spécifique type DIU (structuration par le/la titulaire de la chaire) 3. Implication dans le montage de projets de recherche clinique 4. Participation au montage de projets européens ou internationaux, ERC
Indicateurs secondaires	
Précoce (<2 ans)	1. Publications dans des revues informatiques et médicales spécialisées 2. Participation aux congrès internationaux des sociétés savantes
Moyen terme (2-5 ans)	1. Intégration dans des réseaux de recherche internationaux 2. Organisation d'un séminaire scientifique international 3. Mise en place de contrats partenariaux