La Rochelle Université
Institut Littoral Urbain Durable Intelligent

HR EXCELLENCE IN RESEARCH

PROPOSITION DE SUJET POUR UN CONTRAT DOCTORAL

Laboratoire : L3i

Titre de la thèse

Vers une Efficacité Pareto-Optimale entre Confidentialité et Performance via l'Intégration de l'Apprentissage Fédéré, de la Blockchain et des LLMs pour la Sécurisation des Contenus Sensibles.

Direction de la thèse *nom du/des directeur·trice·s (grade, HDR) et éventuels co-directeur·trice·s uniquement*

INSCRIRE OBLIGATOIREMENT LE POURCENTAGE DE DIRECTION ENVISAGE

Souhail Bakkali (ECC) – Co-Directeur de Thèse (50%)

Mickael Coustaty (MCF, HDR) – Co-Directeur de Thèse (50%)

Contribution aux thématiques scientifiques du LUDI :

Dans un monde de plus en plus numérique, le partage et la gestion des informations sont devenus un enjeu majeur pour les institutions publiques et privées. Par exemple, les documents administratifs, tels que les factures, contrats, et archives historiques, contiennent souvent des données à caractères personnelles qui ne peuvent, ni ne doivent, être partagées. Pourtant, le traitement automatique de ces données nécessite des méthodes de traitement complexes dont la robustesse n'est pas toujours garantie, afin d'assurer la sécurité et la confidentialité de ces données.

Le développement des technologies de capture de données en temps-réel (type Internet des Objets, largement répandus dans les approches de type smart-cities par exemple) est un enjeu majeur de compétitivité des entreprises françaises, afin de permettre une prise de décision systémique et contextualisé. La nécessité d'exploiter efficacement ces ressources tout en respectant les normes de protection des données devient donc cruciale et amène plusieurs défis autour de la protection et de la confidentialité des données.

Une approche récente en apprentissage machine concerne l'apprentissage fédéré (Federated Learning), afin de préserver la vie privée et de garantir la sécurité des données lors de leur traitement. En exploitant cette technique, les institutions peuvent collaborer pour entraîner des modèles d'apprentissage automatique sans partager directement leurs données sensibles, permettant ainsi de mutualiser les jeux de données nécessaires à l'entrainement des systèmes sans déroger aux contraintes de protection des informations personnelles et confidentielles. L'application de modèles de langage (LLMs) pour des tâches telles que la recherche d'information, l'extraction d'information et l'analyse sémantique permet d'exploiter efficacement de vastes ensembles de contenus, facilitant ainsi l'accès à des connaissances précieuses tout en respectant les normes de sécurité et de gouvernance des données. Cependant, l'utilisation croissante des LLMs soulève des défis importants en matière de préservation des données, notamment le risque que ces modèles mémorisent des informations sensibles, compromettant ainsi la confidentialité des données. Il est essentiel de trouver un équilibre entre l'exploitation des capacités avancées des LLMs pour améliorer l'analyse de contenus et la nécessité de protéger les données sensibles, garantissant ainsi une utilisation éthique et responsable de la technologie. Cette thèse s'inscrit dans la dynamique de transition numérique promue par le LUDI, visant à favoriser la dématérialisation, la valorisation des ressources numériques, et à établir des normes de sécurité et de transparence dans la gestion et la gouvernance des données. L'objectif principal étant de permettre l'utilisation d'approches d'IA pour prendre toujours plus de contexte en compte dans les décisions tout en instaurant un cadre de confiance dans nos pratiques numériques et nos échanges d'informations. Ainsi, plusieurs institutions ou chercheurs pourront collaborer au déploiement d'un modèle d'IA commun sans jamais partager leurs données.

Adéquation avec les priorités territoriales (merci de décrire la contribution du sujet de thèses aux thématiques de la CdA et/ou du Département de la Charente Maritime)

Les stratégies développées dans ce projet pourraient avoir des applications pratiques significatives dans plusieurs secteurs, apportant des bénéfices non seulement aux institutions impliquées mais aussi au développement économique et social du département de la Charente Maritime.

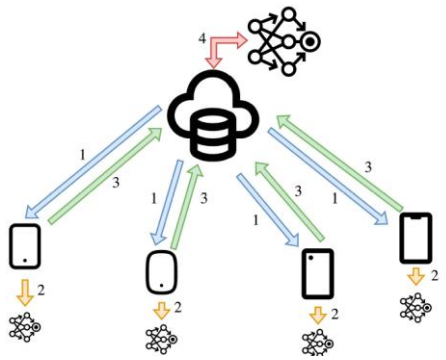
Dans le cadre des bibliothèques universitaires, l'application des technologies d'apprentissage fédéré pourrait révolutionner la manière dont les ressources documentaires sont exploitées. En permettant un accès collaboratif aux données tout en préservant la confidentialité des utilisateurs et des documents, ces outils pourraient faciliter la recherche interdisciplinaire et la mutualisation des connaissances. Par exemple, des systèmes d'extraction d'information pourraient être mis en place pour analyser des catalogues, des archives historiques, et d'autres ressources, permettant aux chercheurs d'accéder à des informations pertinentes sans exposer des données sensibles. Cette démarche renforcerait également la valeur ajoutée des bibliothèques en tant que centres de savoir, soutenant l'innovation et la recherche locale.

Les entreprises cherchant à extraire des informations de contenus variés, tels que des factures, des contrats, ou des rapports d'activité, pourraient grandement bénéficier de ces technologies avancées. Par exemple, l'utilisation de LLMs et d'apprentissage fédéré pour l'automatisation de l'extraction de données peut réduire considérablement le temps et les coûts associés au traitement des contenus. En outre, ces outils permettraient d'améliorer la précision des analyses tout en respectant les normes de confidentialité, renforçant ainsi la confiance des clients et des partenaires. Pour le département, cela pourrait stimuler l'essor d'entreprises innovantes et augmenter l'attractivité du territoire pour des investissements dans le domaine numérique.

L'intégration de la blockchain dans le FL favorise l'émergence de nouveaux modèles commerciaux basés sur la collaboration décentralisée. Des plateformes pourraient voir le jour pour faciliter les échanges et la coopération entre les entreprises, les administrations, et les professionnels de l'assurance ou de la santé, en exploitant les avantages de la blockchain et de l'apprentissage fédéré. De tels modèles commerciaux permettront aux entreprises de partager les bénéfices de la collaboration tout en réduisant les coûts d'exploitation. Ces modèles permettent également de diminuer les dépenses juridiques et administratives associées à la gestion des litiges, tout en réduisant les pertes financières dues aux fraudes. Pour les secteurs comme les assurances, cela permettra de collaborer plus sereinement et à moindre coût.

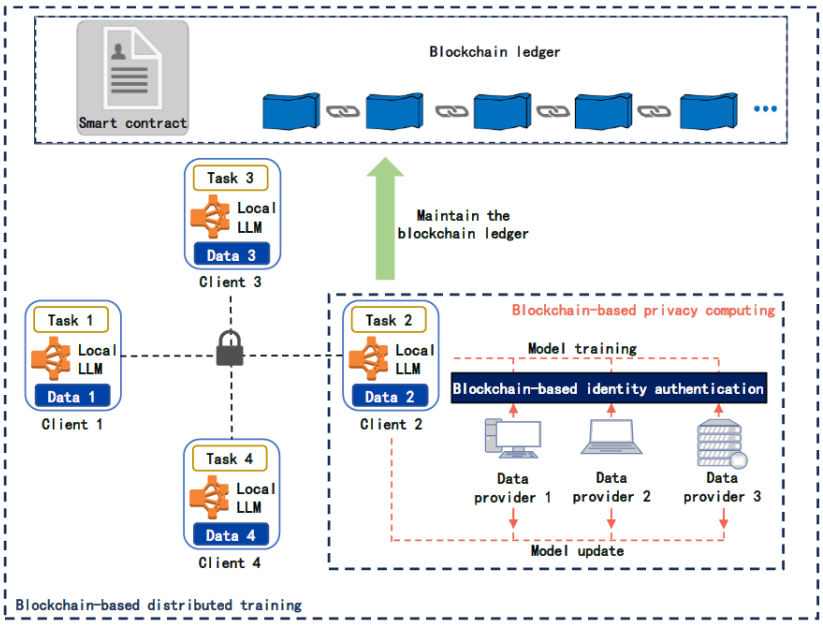
Descriptif du sujet *(enjeux scientifiques, applicatifs, sociétaux…)*

L'intégration de l'apprentissage fédéré (FL) et de la blockchain dans l'analyse des contenus via des modèles de langage de grande taille (LLMs) offre une approche novatrice pour traiter des ensembles de données sensibles tout en préservant la confidentialité. En utilisant l'apprentissage fédéré, les institutions peuvent collaborer pour entraîner des modèles d'apprentissage automatique sans partager directement leurs données, ce qui est essentiel dans des contextes où la protection des informations personnelles est primordiale. Cette collaboration est renforcée par la blockchain, qui permet de créer un cadre décentralisé et sécurisé où chaque interaction et mise à jour des modèles est enregistrée dans un registre immuable. Ainsi, la transparence et la traçabilité des opérations sont assurées, réduisant considérablement le risque d'attaques telles que l'empoisonnement des données ou les intrusions malveillantes.



Bien que la blockchain apporte des avantages en matière de sécurité et de transparence, elle présente aussi des défis en termes de coût computationnel. L'ajout de la blockchain peut donc amplifier les goulots d'étranglement en communication et en calcul déjà présents dans le FL, en particulier si des modèles lourds (LLMs) sont utilisés. Une attention particulière doit être portée à la gestion de l'efficacité pour éviter des surcharges qui ralentiraient le processus d'apprentissage.

Grâce à l'apprentissage fédéré, ces modèles peuvent être formés sur des données provenant de plusieurs sources tout en maintenant ces données localement, minimisant ainsi le risque de fuite d'informations sensibles. La blockchain renforce cette sécurité en vérifiant l'intégrité des mises à jour de modèle, ce qui permet de garantir que les informations traitées et les résultats générés sont authentiques et fiables. En outre, l'utilisation de techniques telles que la confidentialité différentielle lors de l'entraînement des LLMs permet de limiter le risque de mémorisation d'informations spécifiques, garantissant ainsi que les sorties du modèle ne révèlent pas de données sensibles.



Ce cadre combiné d'apprentissage fédéré, de la blockchain et des LLMs offre une solution puissante pour l'analyse des de contenus, facilitant un accès sûr et efficace à des connaissances précieuses tout en répondant aux normes réglementaires, telles que celles imposées par le RGPD. En intégrant ces technologies, cette recherche ambitionne de créer des systèmes qui non seulement optimisent l'interaction avec les ressources numériques, mais aussi renforcent la confiance des utilisateurs dans la gestion de leurs données.

Contexte partenarial *(cotutelle internationale, EU-CONEXUS, partenariat avec un autre laboratoire, une entreprise…)*

La thèse s'effectuera en partenariat avec l'entreprise Siren, dont le thème de cette thèse est en lien avec les activités de l'établissement partenaire. La thèse sera co-dirigée par Adam Jatowt. La possibilité d'une cotutelle internationale est à l'étude.

Impacts *(scientifiques, technologiques, socio-économiques, environnementaux, sociétaux…)*

- **Impacts scientifiques :** L'intégration de la blockchain au FL pousse les frontières de la décentralisation, permettant un cadre plus fiable et plus sécurisé pour l'analyse des données sensibles. Cette approche peut transformer la manière dont les systèmes d'intelligence artificielle et d'apprentissage automatique sont conçus, augmentant la robustesse sans sacrifier la confidentialité. Cette intégration permet également de renforcer la transparence des flux de données et des actions, facilitant ainsi les processus d'audit des contenus sensibles. Cela contribue également au développement de normes sur la traçabilité et la transparence des systèmes d'apprentissage automatique. Le sujet de thèse permet également de développer des mécanismes plus résistants visant à éliminer les différents types d'attaques dans les cadres traditionnels, tels que l'empoisonnement des données ou toute autre forme d'intrusion malveillante. Cette résilience est essentielle dans l'analyse de contenus, où l'intégrité des données est essentielle.

- **Impacts technologiques :** La thèse vise à utiliser la technologie blockchain pour développer des cadres robustes pour l'apprentissage fédéré. Ainsi, le système doit enregistrer chaque interaction ou mise à jour du modèle dans un registre distribué immuable. Cela rend techniquement beaucoup plus difficile pour les attaquants de corrompre les données du modèle. La thèse vise également à explorer des méthodes émergentes telles que le chiffrement homomorphique pour garantir la confidentialité sans compromettre l'efficacité et la précision des modèles. Cela permet de construire des systèmes plus sûrs, adaptés aux exigences strictes de confidentialité dans ces domaines. Enfin, la thèse vise à fournir aux institutions des cadres à utiliser avec

des normes communes pour assurer une gestion décentralisée des contenus numériques. Cela permet de créer des réseaux collaboratifs où chaque participant peut partager des modèles ou des mises à jour sans compromettre la confidentialité des données.

- **Impacts socio-économiques :** L'un des effets socio-économiques les plus significatifs est la protection accrue des données personnelles, notamment dans les secteurs sensibles comme la santé, le droit et l'administration publique. L'utilisation de la blockchain et du FL permet aux institutions de mieux contrôler leurs données, cela renforce la confiance des citoyens et clients envers ces institutions. Sur le plan économique, cette thèse pourrait stimuler l'apparition de nouvelles entreprises et plateformes qui exploitent les avantages du FL et de la blockchain pour fournir des services plus sécurisés, accessibles et optimisés pour le traitement des données sensibles. De plus, grâce à la décentralisation, chaque entité peut utiliser ses propres ressources locales pour traiter les données, tout en bénéficiant des résultats collaboratifs sans avoir besoin d'investir massivement dans des infrastructures centralisées. Cela pourrait réduit les coûts opérationnels pour chaque institution participante.

Programme de travail du doctorant *(tâches confiées au doctorant)*

Le doctorant conduira une revue systématique de la littérature sur l'application du FL dans l'analyse de contenus numériques, en mettant particulièrement l'accent sur les problèmes du FL et les solutions proposées jusqu'à présent. La revue devra examiner les articles récents publiés dans des journaux et conférences pertinentes, qui seront systématiquement analysés et comparés sous plusieurs angles (confidentialité, sécurité, convergence des modèles, coûts de calcul et de communication). À la suite de cette enquête, le doctorant identifiera les défis les plus marquants dans ce contexte.

Mois 1-3 : Phase de Préparation

- **Revue de la littérature :**

Phase 1 : Apprentissage fédéré (FL) et Blockchain pour la sécurité : Le doctorant effectuera une analyse systématique des études existantes sur l'intégration du FL et de la blockchain dans des systèmes sécurisés. Cela inclut l'étude des méthodes pour renforcer la sécurité et garantir la confidentialité dans des systèmes distribués. Le doctorant évaluera également les solutions proposées pour surmonter les limites du FL en termes de communication et de calcul, et les défis liés à l'intégration de la blockchain (notamment les coûts additionnels en ressources et la complexité).

Phase 2 : Modèles de langage de grande taille (LLMs) et Confidentialité des Données : Le doctorant examinera la littérature sur l'utilisation des LLMs pour l'analyse des contenus sensibles, tout en explorant des stratégies pour atténuer les risques liés à la mémorisation des données personnelles par ces modèles. L'accent sera mis sur les techniques de réduction des coûts en utilisant des modèles légers (lightweight models) et les LLMs comme agents autonomes, tout en garantissant la confidentialité des données.

- **Affinement du sujet de thèse :**

Le doctorant, en collaboration avec les superviseurs, affinera les objectifs de la thèse en fonction des découvertes faites lors de la revue de la littérature. Des réunions avec des parties prenantes permettront de définir des priorités claires sur l'intégration de la sécurité (FL et blockchain) et les LLMs.

- **Planification méthodologique :**

Phase 1 : FL et Blockchain : Le doctorant élaborera une méthodologie précise pour l'expérimentation avec FL et blockchain dans des environnements sécurisés. Cela inclura la définition des scénarios de test, la collecte de données et les configurations de blockchain pour optimiser la sécurité et réduire les risques d'attaques (empoisonnement des modèles, fuite de données...).

Phase 2 : LLMs et Confidentialité : La méthodologie d'expérimentation sur les LLMs sera définie. Des techniques d'entraînement sécurisé des modèles seront proposées pour éviter la mémorisation des données à caractère personnel.

Mois 4-12 : Phase de Développement et Expérimentation

Phase 1 : Apprentissage fédéré et Blockchain (Mois 4-6)

- **Implémentation et tests initiaux :**

Le doctorant développera un prototype d'architecture intégrant FL et blockchain pour renforcer la sécurité des échanges de données dans des environnements distribués.

Des expérimentations seront menées pour mesurer l'impact de la blockchain sur la performance et les coûts computationnels du FL, tout en assurant la confidentialité et l'intégrité des données.

- **Analyse préliminaire :**

Le doctorant analysera les premiers résultats obtenus en termes de performance, sécurité et efficacité. Il ajustera la méthodologie si nécessaire pour optimiser la communication entre nœuds FL et blockchain.

Phase 2 : Modèles de Langage de Grande Taille et Confidentialité (Mois 7-12)

- **Implémentation des modèles de langage :**

Le doctorant intégrera des modèles de langage de grande taille (LLMs) dans le cadre d'analyse de contenus numériques, en testant différentes approches pour protéger les données sensibles et réduire la mémoire des informations personnelles par les LLMs. Les techniques de distillation de modèles (lightweight models) seront utilisées pour rendre les LLMs plus efficaces en termes de ressources.

- **Analyse préliminaire :**

Le doctorant analysera les performances des LLMs dans l'extraction d'informations sensibles tout en respectant les exigences de confidentialité, et proposera des ajustements si des problèmes de sécurité ou de mémoire sont détectés.

Mois 13-24 : Phase d'Optimisation et Validation

- **Phase 1 : Sécurisation avec FL et Blockchain (Mois 13-18)**

Le doctorant optimisera l'architecture FL-blockchain pour réduire les coûts computationnels et améliorer l'efficacité tout en maintenant des niveaux de sécurité élevés. Une attention particulière sera portée aux goulots d'étranglement et à l'amélioration de la latence du système. Des simulations de scénarios réels seront réalisées pour valider la solution.

- **Phase 2 : Optimisation des LLMs pour la confidentialité (Mois 19-24)**

Le doctorant ajustera les techniques de protection des données (encryption, confidentialité différentielle, etc.) utilisées dans l'entraînement des LLMs afin de minimiser les risques de fuite d'informations personnelles.
Des scénarios d'application concrets (analyse de factures, documents de santé) seront utilisés pour tester et valider les solutions proposées.

Mois 25-36 : Phase de Finalisation et Rédaction

- **Analyse des résultats et rédaction :**

Le doctorant analysera les résultats finaux de chaque phase (FL + Blockchain et LLMs), rédigera la thèse et préparera des articles scientifiques pour disséminer les résultats dans des conférences et revues internationales.

- **Préparation à la soutenance et défense :**

Simulation de soutenance avec le directeur de thèse et des collègues, révision finale de la thèse et soumission des articles en vue de la publication.

Ce projet de thèse vise à développer une architecture hybride combinant l'apprentissage fédéré, la blockchain et des LLMs pour sécuriser l'analyse des contenus sensibles, tout en assurant un compromis entre performance et confidentialité.

Accompagnement du doctorant / Fonctionnement de la thèse *(accompagnement humain, matériel, financier, en particulier pour la prise en charge du fonctionnement de la thèse et des dépenses associées)*

L'accompagnement du doctorant et le fonctionnement de la thèse sont pris en charge par le laboratoire. Cela inclut la mise à disposition des ressources matérielles nécessaires notamment le matériel informatique et autres équipements appropriés à hauteur de 2 000€. De plus, un financement de 4 000 € est dédié au fonctionnement de la thèse, à la formation doctorale et à la prise en charge de la soutenance de celle-ci. En plus de ces ressources, l'équipe IC du L3i, au travers de ses ressources propres sur ses projets, complétera ces financements avec un budget de 2 000€ pour répondre pleinement aux besoins de cette thèse de doctorat et du doctorant.

Le doctorant bénéficiera d'un encadrement humain solide, principalement assuré par le directeur de thèse et le co-superviseur. Des réunions régulières seront prévues pour discuter de l'avancement du projet, résoudre les problèmes éventuels et orienter le doctorant dans le processus de recherche. Ces interactions favoriseront également le partage d'expériences, la transmission de connaissances et le développement des compétences du doctorant. Des collaborations avec d'autres chercheurs, au sein de l'équipe du projet ou au-delà, sont encouragées pour promouvoir une approche multidisciplinaire et enrichir l'expérience du doctorant.

De plus, le doctorant aura accès à une infrastructure matérielle complète pour mener à bien son projet. Cela inclut l'utilisation des serveurs de calculs puissants équipés des dernières technologies pour les tâches liées à la vision par ordinateur et au traitement automatique du langage naturel au sein du laboratoire L3i. Des espaces de travail dédiés, tels que des salles de réunion, seront mis à disposition pour favoriser un environnement propice à la recherche. L'accès aux bases de données, aux logiciels spécialisés et aux ressources numériques essentielles est également assuré pour permettre au doctorant de mener des analyses approfondies et de produire des résultats de haute qualité. Le fonctionnement de la thèse sera soutenu financièrement pour garantir la réalisation optimale des activités de recherche. Les frais liés à la participation à des conférences, séminaires et ateliers académiques seront couverts, permettant au doctorant de présenter ses travaux, d'échanger avec d'autres chercheurs et de rester informé des développements dans le domaine. Ces ressources financières visent à alléger les contraintes liées à la recherche et à favoriser une contribution significative du doctorant au projet.

En outre, l'accompagnement du doctorant dans le cadre du projet est conçu de manière holistique, visant à garantir un soutien humain, matériel et financier robuste pour favoriser le succès de la thèse et la réalisation des objectifs de recherche.