



## Sujet de thèse

# Génération de $\mu$ -processeur spécifiques par IA générative

**Mots Clés:** Méthodologie de conception, IA générative pour la conception, Architecture des processeurs, jeu d'instructions RISC-V

## 1 Introduction et contexte

Le principal atout de l'architecture de jeu d'instructions RISC-V est qu'elle facilite l'émergence de nouveaux acteurs dans le monde des processeurs. En effet, le fait que cette architecture soit ouverte et libre de droits permet une personnalisation et/ou une déclinaison sans frais de licence. Cette diversité devrait favoriser l'innovation dans la définition de nouvelles architectures de processeurs. Ainsi, ajouter un accélérateur matériel à une architecture de processeur RISC-V présente plusieurs avantages significatifs, en particulier lorsqu'il s'agit d'améliorer les performances pour des tâches spécifiques, tout en maintenant une certaine flexibilité. Les principaux intérêts de cette approche sont :

- Amélioration des performances d'un processeur généraliste pour des opérations très spécifiques ;
- Transfert de l'exécution d'opérations critiques pour que le processeur se concentre sur des tâches plus généralistes ou sur le contrôle de l'exécution globale d'un programme ;
- Personnalisation des architectures de processeurs à des domaines applicatifs ;
- Maîtrise du débit, de la latence et de la consommation du processeur RISC-V.

Dans ce contexte, un des enjeux est la mise à disposition auprès de la communauté de nouveaux outils logiciels d'aide à la conception de processeurs. Ces derniers doivent à la fois faciliter la collaboration entre concepteurs et accélérer le développement de nouvelles déclinaisons de processeur.

L'exploitation du potentiel de l'IA générative dans la définition de nouvelles méthodologies de conception [1] et des outils associés fait l'objet de nombreux travaux de recherche récents. Parmi ces derniers, nous pouvons citer la génération automatique de placement/routage de circuit intégré en utilisant de l'apprentissage par renforcement [2] ou l'utilisation d'un grand modèle de langage (LLM) pour l'exploration de l'espace de solutions architecturales dédiées [3]. Au niveau algorithmique, il existe par exemple des travaux [4] visant à identifier des algorithmes efficaces de multiplication matricielle à l'aide de l'apprentissage par renforcement profond. Plus spécifiquement concernant les processeurs, des premiers travaux [5] ont examiné la faisabilité de l'utilisation de modèles de langage tels que ChatGPT pour générer du code VHDL décrivant un cœur de processeur RISC. Ils mettent en exergue les atouts et les limitations actuelles de ce type d'approche. Par ailleurs, une étude présentée dans [6] explore les défis et les opportunités rencontrés lors de l'utilisation de LLM pour générer des vecteurs de test afin à la fois de faciliter la testabilité et de garantir une couverture de test exhaustive lors de la validation d'architectures de processeur.

## 2 Objectifs scientifiques

Dès lors, exploiter un grand modèle de langage LLM afin d’automatiser la personnalisation d’une architecture de processeur RISC-V nous semble une approche innovante qui se doit d’être explorée. En effet, des LLM peuvent aider à identifier et à intégrer des instructions spécifiques pour des domaines d’application tels que l’intelligence artificielle, le traitement d’image, les communications numériques ou la cryptographie. En interprétant les besoins des utilisateurs finaux ou des développeurs d’applications, un LLM peut proposer des extensions d’instructions sur mesure qui répondent aux exigences particulières des domaines cibles. De plus, il peut faciliter l’automatisation de la co-conception du matériel et du logiciel qu’implique la mise en œuvre d’une déclinaison particulière d’une architecture de processeur RISC-V.

## 3 Profil du/de la candidat(e)

Le/la candidat(e) doit être titulaire d’un master (ou équivalent) et doit être capable de démontrer des connaissances théoriques et pratiques dans les domaines suivants :

- Connaissance de l’architecture des processeurs
- Connaissance en IA générative et grand modèle de langage
- Notions de langage de description matériel (VHDL, SystemVerilog)

## 4 Contexte de la thèse

Ce projet de thèse se situe dans le cadre de l’action CADaBRIA du programme PEPR PHOENIX. L’action CADaBRIA se situe au croisement entre les domaines de l’informatique “architectures des machines” et l’apprentissage automatique. Le/la doctorant(e) sera pleinement intégré(e) (discussions, réunions, séminaires) au sein du projet et collaborera avec les autres chercheurs et doctorants/étudiants de master participants. Il/elle pourra co-encadrer des stagiaires. Des opportunités d’enseignement (rémunérées en plus de son salaire) pourront lui être proposées s’il/elle le souhaite.

Ce projet est une collaboration entre les équipes ASIC du laboratoire CNRS IETR et CSN du laboratoire CNRS IMS. L’équipe ASIC fera bénéficier le/la doctorant(e) de ses compétences en conception de circuits et méthodologie de conception. Quant à l’équipe CSN, elle sera bénéfique à l’étude en raison de son expertise en conception sur FPGA/ASIC de processeurs flexibles à base d’ISA RISC-V. La personne recrutée sera hébergée sur le site de Nantes du laboratoire IETR. Elle sera également amenée à séjourner au laboratoire IMS durant des périodes de la thèse.

- **Laboratoires de recherche** : IETR (UMR CNRS 6164), IMS (UMR CNRS 5218)
- **Université** : Nantes Université
- **Lieu** : IETR/Polytech, Nantes, FRANCE
- **Date de début prévue** : Septembre ou octobre 2026 (durée de 3 ans)

## 5 Candidature

Les candidatures doivent être soumises exclusivement via la plateforme AMETHIS. Aucune candidature reçue en dehors du processus AMETHIS ne sera considérée.

Pour plus d’informations, veuillez contacter :

- **Pr. Sébastien Pillement**, [sebastien.pillement@univ-nantes.fr](mailto:sebastien.pillement@univ-nantes.fr) ,
- **Pr. Christophe Jégo**, [christophe.jego@ims-bordeaux.fr](mailto:christophe.jego@ims-bordeaux.fr) .

Les candidatures seront examinées au fur et à mesure de leur réception. Des auditions seront organisées à l'issue de la période de candidature. Les candidatures ne répondant pas aux qualifications ne seront pas retenues pour les auditions.

L'intégration au laboratoire IETR nécessitant l'accès à une zone sécurisée restreinte, nous vous demandons de prendre en compte le temps nécessaire à l'examen de votre profil si vous souhaitez postuler. En effet, le processus en question peut prendre jusqu'à deux mois.

## References

- [1] Zhuolun He et al. "Large language models for eda: Future or mirage?" In: *ACM Transactions on Design Automation of Electronic Systems* 30.6 (2025), pp. 1–53.
- [2] Azalia Mirhoseini et al. "A graph placement methodology for fast chip design". In: *Nature* 594.7862 (2021), pp. 207–212.
- [3] Kaiyan Chang et al. "Chippgpt: How far are we from natural language hardware design". In: *arXiv preprint arXiv:2305.14019* (2023).
- [4] Alhussein Fawzi et al. "Discovering faster matrix multiplication algorithms with reinforcement learning". In: *Nature* 610.7930 (2022), pp. 47–53.
- [5] Shadeeb Hossain et al. "Using llm such as chatgpt for designing and implementing a risc processor: Execution, challenges and limitations". In: *arXiv preprint arXiv:2401.10364* (2024).
- [6] Chao Xiao et al. "Llm-based processor verification: A case study for neuromorphic processor". In: *2024 Design, Automation & Test in Europe Conference & Exhibition (DATE)*. IEEE. 2024, pp. 1–6.