

The curiosity-driven platform: "Kids Ask"

Curiosity modelling

- Stimulating curiosity**
 - External stimuli
 - Awareness of a missing information
- Sustaining curiosity**
 - Elaboration of the curious thinking
 - Curious seeking of the information

Enciding curiosity

Opportunities to report uncertainty

Training curious question-asking

Encouraging GA-guided autonomous explorations

The experimental design

Group 1: Hand-crafted incentive agent

Utterance 1: You can explore this text more by asking curious questions.

Utterance 2: Here, I can help you think of one: An interesting question can start with "How".

Utterance 3: The answer to this curious question is the following: Can you find it?

The agent renders the checkbox "The Big Bang lasted about a billionth of a second."

Group 2: Automated incentive agent

GPT-3

Utterance 1: You can explore this text more by asking curious questions.

Utterance 2: Here, I can help you think of one: An interesting question can start with "How".

Utterance 3: The answer to this curious question is the following: Can you find it?

The agent renders the checkbox "The Big Bang lasted about a billionth of a second."

Group 3: Automated open agent

GPT-3

Utterance 1: Here are some important key words I could find about this text. You can choose one or several to help you think of a curious question.

The agent renders the checkbox "Explation" and "Dig Deeper".

Utterance 2: I propose to start your question with "but" but you can choose to use another questioning word if you want.

Utterance 3: As a reminder, here is a list of other questioning words that you can use: How, why, Who, ...

Co-supervision: [Pierre-Yves Oudeyer](#) and H    e Sauz    n (Inria), [Edith Law](#) (Univ. Waterloo)

Host: Inria Bordeaux, Flowers project-team (<https://flowers.inria.fr>), in the context of the CuriousTech associate team between Inria and Univ. Waterloo (<https://flowers.inria.fr/curioustech-associate-team>)

Location: Inria Bordeaux (with visits to Univ. Waterloo)

Program/funding: DRI Inria

Duration: 24 months (starting oct. 2024)

How to apply: contact pierre-yves.oudeyer@inria.fr, helene.sauzeon@inria.fr and edithlaw@uwaterloo.ca with CV, letter of motivation and 2 letters of recommendation (this can be sent later), and **using the [application] tag in the email object field. In addition, the application has to be submitted on <https://jobs.inria.fr/public/classic/fr/offres/2024-07626> before 30th may.**

Eligibility: For the candidates who obtained their Ph.D. in the Northern hemisphere, the date of the Ph.D. defense shall be later than September,1 2022; in the Southern hemisphere, later than April,1 2022.

About the Inria centre of the University of Bordeaux and the CuriousTech associate team between Inria and Univ. Waterloo

Inria is the French national research institute for digital science and technology. World-class research, technological innovation and entrepreneurial risk are its DNA. As a technological institute, Inria supports the diversity of innovation pathways: from open source software publishing to the creation of technological startups (Deeptech).

Inria has been present in the Nouvelle-Aquitaine region for more than 20 years thanks to the [Inria center at the University of Bordeaux](#). This center today employs more than 260 people who collaborate through 20 project-teams with 180 people from our academic and industrial partners (universities of Bordeaux, of Bordeaux-Montaigne, and of Pau-et-Pays-de-l'Adour, Bordeaux INP, ENSTA Paris, CNRS, Inserm, INRAE, TotalEnergies and Naval Group).

One of the three main thrusts of its scientific strategy is "**Machine learning and AI**", developed in a unique way in Bordeaux in conjunction with cognitive science, as exemplified by the [FLOWERS project team](#) led by Pierre-Yves Oudeyer. FLOWERS aims to develop the foundations for a new approach of AI and autonomous learning based on the modeling of learning and cognitive development in children, in particular the **mechanisms of curiosity**. This new approach to human-inspired AI naturally finds its ideal application in **educational technologies**.

Working closely with the University of Waterloo on this application area, the international associate team called **CuriousTECH**, was created in January 2023. The [curiousTECH team](#) aims to study **how new educational technologies, using both curiosity-related models and artificial intelligence techniques such as large language models, can personalize learning sequences for each individual, maximizing curiosity and learning efficiency in real world contexts**. This collaboration has already led Hélène Sauzéon, Edith Law and Pierre-Yves Oudeyer to achieve various projects on AI technologies that aim to support curiosity-driven learning, several of which were tested in real world conditions in classrooms thanks to the support from Rectorat and Académoe de Bordeaux with whom we have a partnership. The results were already published in several joint publications (Alaimi et al., 2020; Abdelghani et al. 2023a, 2022), and will form the basis of this postdoc project.

Context

Every year Inria International Relations Department has a few postdoctoral positions in order to support Inria international collaborations. The postdoctoral contract will have a **duration of 12 to 24 months**. The default **start date** is **November 1st, 2024** and **not later than January, 1st 2025**. The **postdoctoral fellow will contribute to the CuriousTech team** and be recruited by the **Inria Centre of the University of Bordeaux in dialog with the University of Waterloo** (Canada) (please note that the postdoctoral fellow has to start his/her contract being in France and that the visits have to respect Inria rules for missions).

Scientific project

Title of the proposal: Design and evaluation of Large-Language Models (LLMs) based conversational agents for fostering curiosity-driven learning in children.

In a constantly changing world full of uncertainties, one way to adapt to unforeseen circumstances is by harnessing lifelong learning driven by curiosity - the desire to acquire information about the world. The nascent field of curiosity research has identified bidirectional links between curiosity and learning: that is, curiosity enhances learning, and vice versa, awareness of one's own learning progress, or metacognition¹, reinforces curiosity (Gottlieb & Oudeyer, 2018; Ten et al., 2021). In this post-doctoral project, we propose an interdisciplinary approach that leverages and will extend recent breakthroughs in understanding how metacognition contributes to the development of curiosity-based learning, aiming to develop novel technology-based educational interventions that could help children develop their curiosity.

Recently, we started to explore how generative AI such as Large Language Models (LLMs) can be used towards positive learning outcomes in educational settings, e.g. through implementing conversational

¹ Classically defined as the ability to reflect on, evaluate and control one's thoughts and mental capabilities, e.g. knowing that one does not know something

agents that train children to ask curiosity-driven learning questions (Abdelghani et al., 2023b, 2023c). However, a limit of this work was that LLMs were used to automate only a small part of the behaviour of conversational agents, for only one form of meta-cognitive training. **The goal of this postdoc will be to scale up this approach by designing and evaluating full fledged LLM-based conversational agents that train children’s curiosity and meta-cognition in a diversity of dimensions. In short, we will aim to set the stage for a generative AI-based tutoring system aimed to train curiosity.**

Recently, a new theoretical framework (Abdelghani et al., 2023b) distinguishes four specific metacognitive skills for triggering and maintaining curiosity: 1) the ability to be aware of missing information in one’s knowledge, 2) the ability to predict the expected learning gain that will result from seeking to resolve a specific problem, 3) the ability to seek the missing information using the appropriate strategy and adapting it to the identified learning goal and finally, 4) the ability to assess the result of the curiosity-driven process and decide of subsequent actions.

Taking inspiration from this new framework, and leveraging LLMs for building verbal incentives to curiosity state (Abdelghani et al., 2023a), we want to develop **conversational agents capable of producing a continuous dialogue with the student, and in particular generate meta-cognitive prompts for children, facilitating the 4 metacognitive stages of a cycle of curiosity-based learning.** This means agents capable of helping students to identify their desired learning objectives (their information gaps) and whether or not they are achievable, then to implement a resolution strategy and evaluate its effectiveness. How could we scale this process up, leveraging the structure of verbal incentives designed by human teachers and automating the production of large quantities of variations of these incentives? The postdoc will study which prompting, in-context learning strategies, and cognitive architecture could enable to design LLM conversational agents that can drive children in diverse forms of conversations with high pedagogical value for fostering curiosity-driven learning. In particular, we will test a library entailing context-specific prompts and generic (context free) prompts for driving respectively in-depth and in-width strategies of curiosity driven learning (Ten et al., 2022).

Series of experiments will be done amongst children for assessing the effectiveness and the efficiency of LLMs-based conversational strategies across the 4-steps of curiosity-learning cycle. To conduct these experiments, we will leverage our ongoing agreements with Académie de Bordeaux and a network of primary and middle schools. We will also aim to study how the generated datasets of metacognitive prompts would be of sufficient quality to train much smaller LLMs, which could have a great potential as this may enable embedding in educational apps and real time personalization of exercise design by teachers. Here, we will follow methodologies used in recent papers studying the impact of high-quality datasets to train high-quality small size LLMs (e.g. “Textbooks are all you need” (Gunasekar, 2023) or the TinyStories project, Eldan and Li, 2023).

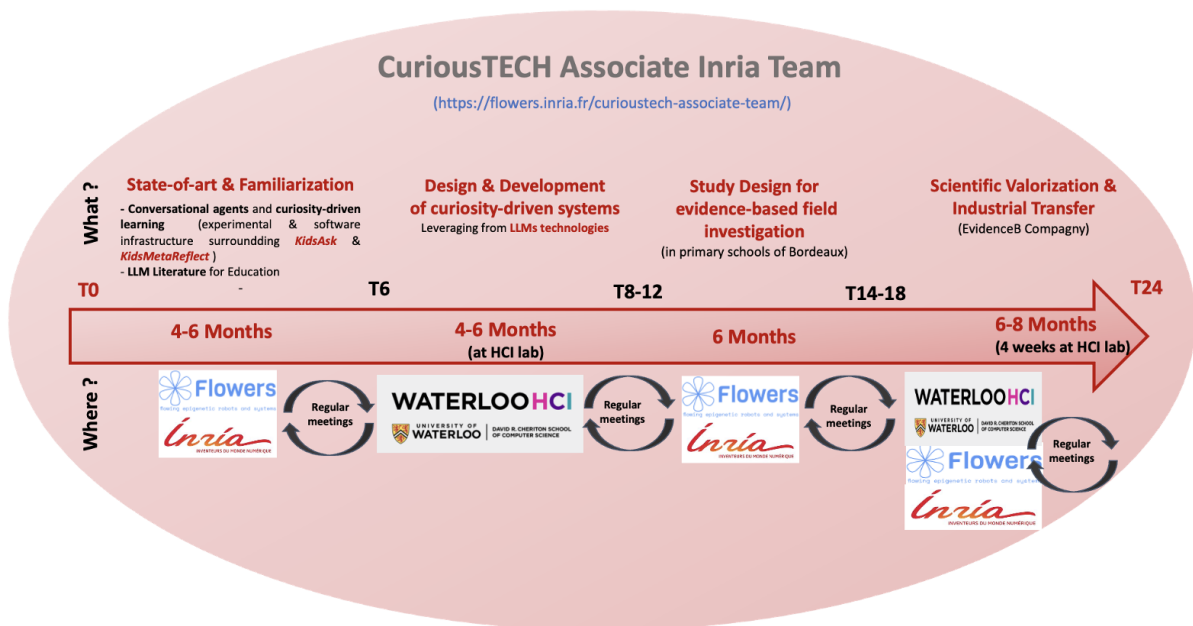
This project will involve a collaboration with Edith Law from the HCI Lab of David R. Cheriton School of Computer Science (University of Waterloo), and world specialist of conversational agents for educational purposes. In particular, the project will leverage the expertise of E. Law’s lab on designing educational technologies that leverage LLMs and HCI design methods, such as in the curiosity notebook approach used in projects studying learning by teaching scenarios.

Also, there will be opportunities in this project to interact with the [EvidenceB](#) company, developing educational technologies leveraging cognitive science and artificial intelligence, and with whom we have several ongoing collaborations. Its educational platform aims to personalize educational contents in order to maximize both learning efficiency and intrinsic motivation in children.

Main Activities

The postdoc will start at Inria Bordeaux by familiarization with the conversational agents, the experimental and software infrastructure used in the *KidsAsk* experiment, as well as the results of the *KidsMetaReflect* experiment. She/he will also familiarize with the recent results studying what are the required characteristics of metacognitive strategies needed to produce robust and generalizable curiosity-driven learning. This will enable to formalize precisely the context and the semantico-syntactic properties of prompts, to address a technical challenge consisting in fine tuning the prompting methods of LLMs for driving the 4 steps of curiosity cycle. In parallel, the candidate will explore various techniques that could be used to implement pedagogically aligned conversational agents using state-of-the-art LLMs, and through interaction with members of the Flowers team doing research on LLMs. Then, the postdoc will visit Edith Law's lab in university of Waterloo (for several weeks) in order to work on HCI dimensions of the design of conversational agents in an educational setting, and leveraging the expertise of E. Law's lab. After the candidate comes back to Inria Bordeaux, she/he will then design an experimental protocol, including appropriate conversational agents and their parameterization, and experimental psychology methods to assess the efficiency, tested with children from primary schools, run these experiments, analyze the results and write a scientific paper(s) describing the work. This work will benefit from various potential collaborations with PhD students and interns both at Inria Flowers and in Edith Law's lab.

Workplan



Timeline of main steps of Postdoctoral project and planned stays at HCI Lab (Waterloo) in the context of CurioTech Associate Inria Team

Assignment

Candidates for postdoctoral positions are recruited **after the end of their Ph.D. or after a first post-doctoral period**: for the candidates who obtained their PhD in the **Northern hemisphere**, the **date of the Ph.D. defense shall be later than September 1, 2022**; in the **Southern hemisphere**, **later than April 1, 2022**. To encourage mobility, the postdoctoral position must take place in a scientific environment that is truly different from the one of the Ph.D. (and, if applicable, from the position

held since the Ph.D.); particular attention is thus paid to French or international candidates who obtained their doctorate abroad.

Skills

Required knowledge and background:

Candidates should have an outstanding expertise in at least one of these areas, and ideally have experience in several of them:

- Experience with LLMs (e.g. through huggingface's transformers library), prompting and/or finetuning will be a plus.
- Digital learning technologies : Educational technologies, Intelligent tutoring systems, e-learning
- Methodologies for assessing educational technologies with users

Other requirements:

- Good skills in programming languages such as python and javascript
- Motivation to work on a project that combines LLM, HCI, cognitive sciences and user studies

References

- Abdelghani, R., Law, E., Desvaux, C., Oudeyer, P. Y., & Sauzéon, H. (2023a).** Interactive environments for training children's curiosity through the practice of metacognitive skills: a pilot study. In *Proceedings of the 22nd Annual ACM Interaction Design and Children Conference*. 495-501.
- Abdelghani, R., Oudeyer, P. Y., Law, E., de Vulpillières, C., & Sauzéon, H. (2022).** Conversational agents for fostering curiosity-driven learning in children. *Int. J. Hum. Comput.* 167, 102887.
- Abdelghani, R., Wang, Y. H., Yuan, X., Wang, T., Lucas, P., Sauzéon, H., & Oudeyer, P. Y. (2023b).** GPT-3-driven pedagogical agents to train children's curious question-asking skills. *International Journal of Artificial Intelligence in Education*, 1-36.
- Abdelghani, R., Sauzéon, H., & Oudeyer, P. Y. (2023c).** Generative AI in the Classroom: Can Students Remain Active Learners? *Workshop on Generative AI in education, Neurips 2023*.
- Alaimi, Mehdi, et al. "Pedagogical agents for fostering question-asking skills in children." *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. 2020.**
- Eldan, R., & Li, Y. (2023).** TinyStories: How Small Can Language Models Be and Still Speak Coherent English?. *arXiv preprint arXiv:2305.07759*.
- Gunasekar, S., Zhang, Y., Aneja, J., Mendes, C. C. T., Del Giorno, A., Gopi, S., ... & Li, Y. (2023).** Textbooks Are All You Need. *arXiv preprint arXiv:2306.11644*.
- Gottlieb, J., & Oudeyer, P. Y. (2018).** Towards a neuroscience of active sampling and curiosity. *Nature Reviews Neuroscience*, 19(12), 758-770.
- Ten, A., Kaushik, P., Oudeyer, P. Y., & Gottlieb, J. (2021).** Humans monitor learning progress in curiosity-driven exploration. *Nat. Commun.* 12(1), 5972.
- Ten, A., Oudeyer, P. Y., & Moulin-Frier, C. (2022).** Curiosity-driven exploration. *The Drive for Knowledge: The Science of Human Information Seeking*, 53.
- Xiao, Z., Yuan, X., Liao, Q. V., Abdelghani, R., & Oudeyer, P. Y. (2023).** Supporting Qualitative Analysis with Large Language Models: Combining Codebook with GPT-3 for Deductive Coding. In *Companion Proceedings of the 28th International Conference on Intelligent User Interfaces.*, 75-78.

Deadline for application : May 30th, 2024.