

RESEARCH CENTRE

**Inria Centre at the University of  
Bordeaux**

IN PARTNERSHIP WITH:

**Ecole nationale supérieure des techniques  
avancées**

2024

ACTIVITY REPORT

Team

FLOWERS

## **Flowing Epigenetic Robots and Systems**

Inria teams are typically groups of researchers working on the definition of a common project, and objectives, with the goal to arrive at the creation of a project-team. Such project-teams may include other partners (universities or research institutions)

**DOMAIN**

**Perception, Cognition and Interaction**

**THEME**

**Robotics and Smart environments**

*Inria*

# Contents

<b>Team FLOWERS</b>	<b>1</b>
<b>1 Team members, visitors, external collaborators</b>	<b>3</b>
<b>2 Overall objectives</b>	<b>4</b>
<b>3 Research program</b>	<b>6</b>
<b>4 Application domains</b>	<b>10</b>
<b>5 Social and environmental responsibility</b>	<b>11</b>
5.1 Footprint of research activities	11
5.2 Impact of research results	11
<b>6 Highlights of the year</b>	<b>13</b>
6.1 Awards	14
<b>7 New software, platforms, open data</b>	<b>15</b>
7.1 New software	15
7.1.1 SocialAI	15
7.1.2 AutoDisc	15
7.1.3 ADTool	15
7.1.4 Kids Ask	16
7.1.5 ToGather	16
7.1.6 mc_training	16
7.1.7 Evolution of adaptation mechanisms in complex environments	17
7.1.8 SAPIENS	18
7.1.9 architect-builder-abig	18
7.1.10 EAGER	19
7.1.11 Flow-Lenia	19
7.1.12 Kidlearn: money game application	19
7.1.13 cognitive-testbattery	20
7.1.14 Sensorimotor-lenia	21
7.1.15 Lamorel	21
7.1.16 GLAM	22
7.1.17 SBMLtoODEjax	22
7.1.18 Vivarium	23
7.1.19 LLM_Culture	23
7.1.20 TelephoneGameLLMs	24
7.1.21 styr	24
7.1.22 transformerXL_PPO_JAX	24
7.1.23 ER-MRL	25
7.2 New platforms	25
7.2.1 ToGather application	25
7.3 Open data	25
<b>8 New results</b>	<b>25</b>
8.1 Open-ended learning and autotelic AI with large language models	26
8.1.1 ACES: Generating a Diversity of Challenging Programming Puzzles with Autotelic Generative Models	26
8.1.2 MAGELLAN: Metacognitive Generalization of Learning Progress for Online RL in LLM agents	27
8.1.3 SAC-GLAM: Improving Online RL for LLM agents with Soft Actor-Critic and Hindsight Relabeling	28

8.1.4	Learning when (not) to use tools with Memory-based In-context learning and online multi-objective RL in LLM agents . . . . .	29
8.1.5	WorldLLM: Improving LLMs' world modelling using Curiosity-driven Theory-making . . . . .	31
8.1.6	Reinforcement Learning for Aligning Large Language Models Agents with Interactive Environments: Quantifying and Mitigating Prompt Overfitting . . . . .	32
8.1.7	Jack of All Trades, Master of Some, a Multi-Purpose Transformer Agent . . . . .	32
8.1.8	LLM-based goal generation for autotelic exploration with goal-conditioned RL . . . . .	33
8.2	Models of cultural evolution in humans and AI systems . . . . .	34
8.2.1	The effect of social network structure on collective innovation . . . . .	34
8.2.2	The SocialAI School: Insights from Developmental Psychology Towards Artificial Socio-Cultural Agents . . . . .	36
8.2.3	Stick to your Role! Stability of Personal Values Expressed in Large Language Models . . . . .	37
8.2.4	Cultural evolution in populations of Large Language Models . . . . .	37
8.2.5	When LLMs Play the Telephone Game: Cumulative Changes and Attractors in Iterated Cultural Transmissions . . . . .	39
8.2.6	Collapse of diversity and quality in the cultural evolution of Large Language models . . . . .	40
8.2.7	Inferring the Phylogeny of Large Language Models and Predicting their Performances in Benchmarks . . . . .	41
8.3	An Eco-Evo-Devo perspective on Artificial Intelligence . . . . .	43
8.3.1	Research perspective: The Ecology of Open-Ended skill Acquisition . . . . .	43
8.3.2	Eco-evolutionary Dynamics of Non-episodic Neuroevolution in Large Multi-agent Environments . . . . .	44
8.3.3	Emergent kin selection of altruistic feeding via non-episodic neuroevolution . . . . .	45
8.3.4	Evolving large populations of adaptive neural agents in ecologically plausible environments . . . . .	47
8.3.5	Cooperative control of environmental extremes by artificial intelligent agents . . . . .	48
8.3.6	Emergence of agriculture in an artificial society of reinforcement learning agents . . . . .	48
8.3.7	Open-ended recipe crafting through meta reinforcement learning . . . . .	49
8.3.8	Evolving Reservoirs for Meta Reinforcement Learning . . . . .	51
8.4	Theories of human curiosity-driven learning . . . . .	52
8.5	Generative AI and educational technologies . . . . .	54
8.5.1	Fostering curiosity and meta-cognition in children using LLM-based conversational agents . . . . .	54
8.5.2	LLMs usages and curiosity-related metacognitive skills in educational setting . . . . .	56
8.5.3	Supporting Qualitative Analysis with Large Language Models: Combining Codebook with GPT-3 for Deductive Coding . . . . .	61
8.6	Curiosity-driven educational technologies . . . . .	61
8.6.1	New digital approaches for studying curiosity-driven learning . . . . .	61
8.6.2	Machine Learning for Adaptive Personalization in Intelligent Tutoring Systems . . . . .	62
8.6.3	Machine learning for adaptive cognitive training . . . . .	65
8.6.4	ToGather : Interactive website to foster collaboration among stakeholders of school inclusion for pupils with neurodevelopmental disorders . . . . .	70
8.6.5	Curious and therefore not overloaded : Study of the links between curiosity and cognitive load in learning mediated by immersive technologies . . . . .	71
8.6.6	Self-determination-driven digital services for supporting aging-in place and well-being: a study of relationships between longitudinal data from smart home and clinical data . . . . .	73
8.7	Curiosity-driven AI for assisted scientific discovery . . . . .	74
8.7.1	Design of an Interactive Software for Automated Discovery in Complex Systems . . . . .	74
8.7.2	Exploration of Gene Regulatory Network Behaviors Using Automated Discovery Tools . . . . .	74
8.7.3	Discovering Sensorimotor Agency in Cellular Automata using Diversity Search . . . . .	76
8.7.4	Flow lenia: Mass conservation for the study of virtual creatures in continuous cellular automata . . . . .	78
8.7.5	Semantic Open-Endedness in Flow-Lenia using Vision Language Models and IMGEP . . . . .	79

<b>9</b>	<b>Bilateral contracts and grants with industry</b>	<b>79</b>
9.1	Bilateral contracts with industry	79
9.2	Bilateral Grants with Fundation	80
<b>10</b>	<b>Partnerships and cooperations</b>	<b>81</b>
10.1	International initiatives	81
10.1.1	Associate Teams in the framework of an Inria International Lab or in the framework of an Inria International Program	81
10.2	International research visitors	82
10.2.1	Visits of international scientists	82
10.2.2	Visits to international teams	82
10.3	European initiatives	82
10.3.1	Horizon Europe	82
10.3.2	Other european programs/initiatives	83
10.4	National initiatives	84
10.4.1	Adaptiv'Math	85
<b>11</b>	<b>Dissemination</b>	<b>85</b>
11.1	Scientific events: organisation	85
11.1.1	Invited talks	87
11.1.2	Leadership within the scientific community	87
11.1.3	Scientific expertise	88
11.1.4	Research administration	88
11.2	Teaching - Supervision - Juries	88
11.2.1	Teaching	88
11.2.2	Supervision	90
11.2.3	Juries	91
11.3	Popularization	92
11.3.1	Specific official responsibilities in science outreach structures	93
11.3.2	Productions (articles, videos, podcasts, serious games, ...)	93
11.3.3	Participation in Live events	93
11.3.4	Others science outreach relevant activities	94
<b>12</b>	<b>Scientific production</b>	<b>95</b>
12.1	Major publications	95
12.2	Publications of the year	98
12.3	Cited publications	102

## Team FLOWERS

*Creation of the Team: 2023 July 01*

### Keywords

#### Computer sciences and digital sciences

- A5.1.1. – Engineering of interactive systems
- A5.1.2. – Evaluation of interactive systems
- A5.1.4. – Brain-computer interfaces, physiological computing
- A5.1.5. – Body-based interfaces
- A5.1.6. – Tangible interfaces
- A5.1.7. – Multimodal interfaces
- A5.3.3. – Pattern recognition
- A5.4.1. – Object recognition
- A5.4.2. – Activity recognition
- A5.7.3. – Speech
- A5.8. – Natural language processing
- A5.10.5. – Robot interaction (with the environment, humans, other robots)
- A5.10.7. – Learning
- A5.10.8. – Cognitive robotics and systems
- A5.11.1. – Human activity analysis and recognition
- A6.3.1. – Inverse problems
- A9.2. – Machine learning
- A9.4. – Natural language processing
- A9.5. – Robotics
- A9.7. – AI algorithmics

#### Other research topics and application domains

- B1.2.1. – Understanding and simulation of the brain and the nervous system
- B1.2.2. – Cognitive science
- B5.6. – Robotic systems
- B5.7. – 3D printing
- B5.8. – Learning and training
- B9. – Society and Knowledge
- B9.1. – Education
- B9.1.1. – E-learning, MOOC
- B9.2. – Art
- B9.2.1. – Music, sound
- B9.2.4. – Theater

B9.6. – Humanities

B9.6.1. – Psychology

B9.6.8. – Linguistics

B9.7. – Knowledge dissemination

# 1 Team members, visitors, external collaborators

## Research Scientists

- Pierre-Yves Oudeyer [Team leader, INRIA, Senior Researcher]
- Clément Moulin-Frier [INRIA, Researcher]
- Hélène Sauzéon [INRIA, Professor Detachement]

## Faculty Member

- Cécile Mazon [UNIV BORDEAUX, Associate Professor]

## Post-Doctoral Fellows

- Olivier Clerc [INRIA, Post-Doctoral Fellow, from Nov 2024]
- Cedric Colas [INRIA, Post-Doctoral Fellow]
- Eleni Nisioti [UNIV COPENHAGUE]
- Marion Pech [INRIA, Post-Doctoral Fellow]
- Maria Teodorescu [INRIA, Post-Doctoral Fellow]

## PhD Students

- Rania Abdelghani [INRIA, from May 2024 until Sep 2024]
- Rania Abdelghani [EVIDENCEB, until May 2024]
- Maxime Adolphe [ONEPOINT, until Apr 2024]
- Timothé BOULET [INRIA, from Dec 2024]
- Timothe Boulet [INRIA, from Dec 2024]
- Thomas Carta [UNIV BORDEAUX]
- Marko Cvjetko [UNIV BORDEAUX, from Oct 2024]
- Marie-Sarah Desvaux [UNIV BORDEAUX]
- Juliette Deyts [UNIV BORDEAUX, from Nov 2024]
- Loris Gaven [INRIA, from Oct 2024]
- Gautier Hamon [INRIA]
- Sina Khajehabdollahi [INRIA, from Jun 2024]
- Grgur Kovac [INRIA]
- Jeremy Perez [UNIV BORDEAUX]
- Matisse Poupard [CATIE, CIFRE]
- Guillaume Pourcel [INRIA, from May 2024 until Nov 2024]
- Julien Pourcel [INRIA]
- Clément Romac [HUGGING FACE SAS, CIFRE]
- Isabeau Saint-Supery [UNIV BORDEAUX, from Nov 2024]
- Isabeau Saint-Supery [UNIV BORDEAUX, until Aug 2024]
- Nicolas Yax [ENS Paris]

### Technical Staff

- Zacharie Bugaud [INRIA, Engineer, from Mar 2024]
- Corentin Leger [INRIA, Engineer]
- Jesse Lin [INRIA]

### Interns and Apprentices

- Timothe Boulet [INRIA, Intern, from May 2024 until Oct 2024]
- Loris Gaven [INRIA, Intern, from Mar 2024 until Sep 2024]
- Alexis Gerard [INRIA, Intern, from Feb 2024 until May 2024]
- Guillaume Levy [INRIA, Intern, from Apr 2024 until Oct 2024]
- Martial Marzloff [INRIA, Intern, from Feb 2024 until Apr 2024]
- Thomas Michel [INRIA, Intern, from Sep 2024 until Sep 2024]
- Thomas Michel [ENS PARIS-SACLAY, Intern, from Apr 2024 until Aug 2024]
- Gaia Molinaro [INRIA, Intern, until Jun 2024]
- Carole Morvan [UNIV BORDEAUX, Intern, from Apr 2024 until Jun 2024]
- Lucas Spooner [UNIV BORDEAUX, from May 2024 until Jun 2024]

### Administrative Assistant

- Nathalie Robin [INRIA]

### Visiting Scientist

- Max Taylor-Davies [UNIV EDIMBOURG, from Jun 2024 until Aug 2024]

### External Collaborators

- Maxime Adolphe [ONEPOINT, from Apr 2024]
- Didier Roy [LearnSyst]

## 2 Overall objectives

*Abstract:* The Flowers project-team studies models of open-ended development and learning. These models are used as tools to help us understand better how children learn, as well as to build machines that learn like children, i.e. developmental artificial intelligence, with applications in educational technologies, assisted scientific discovery, video games, robotics and human-computer interaction.

*Context:* Great advances have been made recently in artificial intelligence concerning the topic of how autonomous agents can learn to act in uncertain and complex environments, thanks to the development of advanced Deep Reinforcement Learning techniques. These advances have for example led to impressive results with AlphaGo [176] or algorithms that learn to play video games from scratch [148, 127]. However, these techniques are still far away from solving the ambitious goal of lifelong autonomous machine learning of repertoires of skills in real-world, large and open environments. They are also very far from the capabilities of human learning and cognition. Indeed, developmental processes allow humans, and especially infants, to continuously acquire novel skills and adapt to their environment over their entire lifetime. They do so autonomously, i.e. through a combination of self-exploration and



linguistic/social interaction with their social peers, sampling their own goals while benefiting from the natural language guidance of their peers, and without the need for an “engineer” to open and retune the brain and the environment specifically for each new task (e.g. for providing a task-specific external reward channel). Furthermore, humans are extremely efficient at learning fast (few interactions with their environment) skills that are very high-dimensional both in perception and action, while being embedded in open changing environments with limited resources of time, energy and computation.

Thus, a major scientific challenge in artificial intelligence and cognitive sciences is to understand how humans and machines can efficiently acquire world models, as well as open and cumulative repertoires of skills over an extended time span. Processes of sensorimotor, cognitive and social development are organized along ordered phases of increasing complexity, and result from the complex interaction between the brain/body with its physical and social environment. Making progress towards these fundamental scientific challenges is also crucial for many downstream applications. Indeed, autonomous lifelong learning capabilities similar to those shown by humans are key requirements for developing virtual or physical agents that need to continuously explore and adapt skills for interacting with new or changing tasks, environments, or people. This is crucial for applications like assistive technologies with non-engineer users, such as robots or virtual agents that need to explore and adapt autonomously to new environments, adapt robustly to potential damages of their body, or help humans to learn or discover new knowledge in education settings, and need to communicate through natural language with human users, grounding the meaning of sentences into their sensorimotor representations.

*The Developmental AI approach:* Human and biological sciences have identified various families of developmental mechanisms that are key to explain how infants can acquire so robustly a wide diversity of skills [129, 147], in spite of the complexity and high-dimensionality of the body [91] and the open-endedness of its potential interactions with the physical and social environment. To advance the fundamental understanding of these mechanisms of development as well as their transposition in machines, the FLOWERS team has been developing an approach called *Developmental artificial intelligence*, leveraging and integrating ideas and techniques from developmental robotics ([189, 138, 95, 155]), Deep (Reinforcement) Learning and developmental psychology. This approach consists in developing computational models that leverage advanced machine learning techniques such as intrinsically motivated Deep Reinforcement Learning, in strong collaboration with developmental psychology and neuroscience. In particular, the team focuses on models of intrinsically motivated learning and exploration (also called curiosity-driven learning), with mechanisms enabling agents to learn to represent and generate their own goals, self-organizing a learning curriculum for efficient learning of world models and skill repertoire under limited resources of time, energy and compute. The team also studies how autonomous learning mechanisms can enable humans and machines to acquire and develop grounded and culturally shared language skills, using neuro-symbolic architectures for learning structured representations and handling systematic compositionality and generalization.

Our fundamental research is organized along three strands:

**Strand 1: Lifelong autonomous learning in machines.** Understanding how developmental mechanisms can be functionally formalized/transposed in machines and explore how they can allow these machines to acquire efficiently open-ended repertoires of skills through self-exploration and social interaction.

**Strand 2: Computational models as tools to understand human development in cognitive sciences.**

The computational modelling of lifelong learning and development mechanisms achieved in the team centrally targets to contribute to our understanding of the processes of sensorimotor, cognitive and social development in humans. In particular, it provides a methodological basis to analyze the dynamics of interactions across learning and inference processes, embodiment and the social environment, allowing to formalize precise hypotheses and later on test them in experimental paradigms with animals and humans. A paradigmatic example of this activity is the Neurocuriosity project achieved in collaboration with the cognitive neuroscience lab of Jacqueline Gottlieb, where theoretical models of the mechanisms of information seeking, active learning and spontaneous exploration have been developed in coordination with experimental evidence and investigation [18, 31].

**Strand 3: Applications.** Beyond leading to new theories and new experimental paradigms to understand

human development in cognitive science, as well as new fundamental approaches to developmental machine learning, the team explores how such models can find applications in robotics, human-computer interaction, multi-agent systems, automated discovery and educational technologies. In robotics, the team studies how artificial curiosity combined with imitation learning can provide essential building blocks allowing robots to acquire multiple tasks through natural interaction with naive human users, for example in the context of assistive robotics. The team also studies how models of curiosity-driven learning can be transposed in algorithms for intelligent tutoring systems, allowing educational software to incrementally and dynamically adapt to the particularities of each human learner, and proposing personalized sequences of teaching activities.

### 3 Research program

Research in artificial intelligence, machine learning and pattern recognition has produced a tremendous amount of results and concepts in the last decades. A blooming number of learning paradigms - supervised, unsupervised, reinforcement, active, associative, symbolic, connectionist, situated, hybrid, distributed learning... - nourished the elaboration of highly sophisticated algorithms for tasks such as visual object recognition, speech recognition, robot walking, grasping or navigation, the prediction of stock prices, the evaluation of risk for insurances, adaptive data routing on the internet, etc... Yet, we are still very far from being able to build machines capable of adapting to the physical and social environment with the flexibility, robustness, and versatility of a one-year-old human child.

Indeed, one striking characteristic of human children is the nearly open-ended diversity of the skills they learn. They not only can improve existing skills, but also continuously learn new ones. If evolution certainly provided them with specific pre-wiring for certain activities such as feeding or visual object tracking, evidence shows that there are also numerous skills that they learn smoothly but could not be “anticipated” by biological evolution, for example learning to drive a tricycle, using an electronic piano toy or using a video game joystick. On the contrary, existing learning machines, and robots in particular, are typically only able to learn a single pre-specified task or a single kind of skill. Once this task is learnt, for example walking with two legs, learning is over. If one wants the robot to learn a second task, for example grasping objects in its visual field, then an engineer needs to re-program manually its learning structures: traditional approaches to task-specific machine/robot learning typically include engineer choices of the relevant sensorimotor channels, specific design of the reward function, choices about when learning begins and ends, and what learning algorithms and associated parameters shall be optimized.

As can be seen, this requires a lot of important choices from the engineer, and one could hardly use the term “autonomous” learning. On the contrary, human children do not learn following anything looking like that process, at least during their very first years. Babies develop and explore the world by themselves, focusing their interest on various activities driven both by internal motives and social guidance from adults who only have a folk understanding of their brains. Adults provide learning opportunities and scaffolding, but eventually young babies always decide for themselves what activity to practice or not. Specific tasks are rarely imposed to them. Yet, they steadily discover and learn how to use their body as well as its relationships with the physical and social environment. Also, the spectrum of skills that they learn continuously expands in an organized manner: they undergo a developmental trajectory in which simple skills are learnt first, and skills of progressively increasing complexity are subsequently learnt.

A link can be made to educational systems where research in several domains have tried to study how to provide a good learning or training experience to learners. This includes the experiences that allow better learning, and in which sequence they must be experienced. This problem is complementary to that of the learner who tries to progress efficiently, and the teacher here has to use as efficiently the limited time and motivational resources of the learner. Several results from psychology [90] and neuroscience [116] have argued that the human brain feels intrinsic pleasure in practicing activities of optimal difficulty or challenge. A teacher must exploit such activities to create positive psychological states of flow [108] for fostering the individual engagement in learning activities. A such view is also relevant for reeducation issues where inter-individual variability, and thus intervention personalization are challenges of the same magnitude as those for education of children.

A grand challenge is thus to be able to build machines that possess this capability to discover, adapt and develop continuously new know-how and new knowledge in unknown and changing environments,

like human children. In 1950, Turing wrote that the child's brain would show us the way to intelligence: "Instead of trying to produce a program to simulate the adult mind, why not rather try to produce one which simulates the child's" [186]. Maybe, in opposition to work in the field of Artificial Intelligence who has focused on mechanisms trying to match the capabilities of "intelligent" human adults such as chess playing or natural language dialogue [123], it is time to take the advice of Turing seriously. This is what a new field, called developmental (or epigenetic) robotics, is trying to achieve [138] [189]. The approach of developmental robotics consists in importing and implementing concepts and mechanisms from developmental psychology [147], cognitive linguistics [107], and developmental cognitive neuroscience [128] where there has been a considerable amount of research and theories to understand and explain how children learn and develop. A number of general principles are underlying this research agenda: embodiment [94] [157], grounding [121], situatedness [180], self-organization [182] [156], enaction [187], and incremental learning [100].

Among the many issues and challenges of developmental robotics, two of them are of paramount importance: exploration mechanisms and mechanisms for abstracting and making sense of initially unknown sensorimotor channels. Indeed, the typical space of sensorimotor skills that can be encountered and learnt by a developmental robot, as those encountered by human infants, is immensely vast and inhomogeneous. With a sufficiently rich environment and multimodal set of sensors and effectors, the space of possible sensorimotor activities is simply too large to be explored exhaustively in any robot's life time: it is impossible to learn all possible skills and represent all conceivable sensory percepts. Moreover, some skills are very basic to learn, some other very complicated, and many of them require the mastery of others in order to be learnt. For example, learning to manipulate a piano toy requires first to know how to move one's hand to reach the piano and how to touch specific parts of the toy with the fingers. And knowing how to move the hand might require to know how to track it visually.

Exploring such a space of skills randomly is bound to fail or result at best on very inefficient learning [158]. Thus, exploration needs to be organized and guided. The approach of epigenetic robotics is to take inspiration from the mechanisms that allow human infants to be progressively guided, i.e. to develop. There are two broad classes of guiding mechanisms which control exploration:

1. **internal guiding mechanisms**, and in particular intrinsic motivation, responsible of spontaneous exploration and curiosity in humans, which is one of the central mechanisms investigated in FLOWERS, and technically amounts to achieve online active self-regulation of the growth of complexity in learning situations;
2. **social learning and guidance**, a learning mechanisms that exploits the knowledge of other agents in the environment and/or that is guided by those same agents. These mechanisms exist in many different forms like emotional reinforcement, stimulus enhancement, social motivation, guidance, feedback or imitation, some of which being also investigated in FLOWERS.

**Internal guiding mechanisms** In infant development, one observes a progressive increase of the complexity of activities with an associated progressive increase of capabilities [147], children do not learn everything at one time: for example, they first learn to roll over, then to crawl and sit, and only when these skills are operational, they begin to learn how to stand. The perceptual system also gradually develops, increasing children perceptual capabilities other time while they engage in activities like throwing or manipulating objects. This make it possible to learn to identify objects in more and more complex situations and to learn more and more of their physical characteristics.

Development is therefore progressive and incremental, and this might be a crucial feature explaining the efficiency with which children explore and learn so fast. Taking inspiration from these observations, some roboticists and researchers in machine learning have argued that learning a given task could be made much easier for a robot if it followed a developmental sequence and "started simple" [88] [113]. However, in these experiments, the developmental sequence was crafted by hand: roboticists manually build simpler versions of a complex task and put the robot successively in versions of the task of increasing complexity. And when they wanted the robot to learn a new task, they had to design a novel reward function.

Thus, there is a need for mechanisms that allow the autonomous control and generation of the developmental trajectory. Psychologists have proposed that intrinsic motivations play a crucial role.

Intrinsic motivations are mechanisms that push humans to explore activities or situations that have intermediate/optimal levels of novelty, cognitive dissonance, or challenge [90] [108] [110]. Further, the exploration of critical role of intrinsic motivation as lever of cognitive development for all and for all ages is today expanded to several fields of research, closest to its original study, special education or cognitive aging, and farther away, neuropsychological clinical research. The role and structure of intrinsic motivation in humans have been made more precise thanks to recent discoveries in neuroscience showing the implication of dopaminergic circuits and in exploration behaviours and curiosity [109] [124] [174]. Based on this, a number of researchers have begun in the past few years to build computational implementation of intrinsic motivation [158] [159] [171] [89] [125] [140] [172]. While initial models were developed for simple simulated worlds, a current challenge is to manage to build intrinsic motivation systems that can efficiently drive exploratory behaviour in high-dimensional unprepared real world robotic sensorimotor spaces [159, 158, 160, 170]. Specific and complex problems are posed by real sensorimotor spaces, in particular due to the fact that they are both high-dimensional as well as (usually) deeply inhomogeneous. As an example for the latter issue, some regions of real sensorimotor spaces are often unlearnable due to inherent stochasticity or difficulty, in which case heuristics based on the incentive to explore zones of maximal unpredictability or uncertainty, which are often used in the field of active learning [103] [122] typically lead to catastrophic results. The issue of high dimensionality does not only concern motor spaces, but also sensory spaces, leading to the problem of correctly identifying, among typically thousands of quantities, those latent variables that have links to behavioral choices. In FLOWERS, we aim at developing intrinsically motivated exploration mechanisms that scale in those spaces, by studying suitable abstraction processes in conjunction with exploration strategies.

**Socially Guided and Interactive Learning** Social guidance is as important as intrinsic motivation in the cognitive development of human babies [147]. There is a vast literature on learning by demonstration in robots where the actions of humans in the environment are recognized and transferred to robots [87]. Most such approaches are completely passive: the human executes actions and the robot learns from the acquired data. Recently, the notion of interactive learning has been introduced in [183, 93], motivated by the various mechanisms that allow humans to socially guide a robot [166]. In an interactive context the steps of self-exploration and social guidance are not separated and a robot learns by self exploration and by receiving extra feedback from the social context [183, 132, 141].

Social guidance is also particularly important for learning to segment and categorize the perceptual space. Indeed, parents interact a lot with infants, for example teaching them to recognize and name objects or characteristics of these objects. Their role is particularly important in directing the infant attention towards objects of interest that will make it possible to simplify at first the perceptual space by pointing out a segment of the environment that can be isolated, named and acted upon. These interactions will then be complemented by the children own experiments on the objects chosen according to intrinsic motivation in order to improve the knowledge of the object, its physical properties and the actions that could be performed with it.

In FLOWERS, we are aiming at including intrinsic motivation system in the self-exploration part thus combining efficient self-learning with social guidance [152, 153]. We also work on developing perceptual capabilities by gradually segmenting the perceptual space and identifying objects and their characteristics through interaction with the user [139] and robots experiments [126]. Another challenge is to allow for more flexible interaction protocols with the user in terms of what type of feedback is provided and how it is provided [137].

Exploration mechanisms are combined with research in the following directions:

**Cumulative learning, reinforcement learning and optimization of autonomous skill learning** FLOWERS develops machine learning algorithms that can allow embodied machines to acquire cumulatively sensorimotor skills. In particular, we develop optimization and reinforcement learning systems which allow robots to discover and learn dictionaries of motor primitives, and then combine them to form higher-level sensorimotor skills.

**Autonomous perceptual and representation learning** In order to harness the complexity of perceptual and motor spaces, as well as to pave the way to higher-level cognitive skills, developmental learning

requires abstraction mechanisms that can infer structural information out of sets of sensorimotor channels whose semantics is unknown, discovering for example the topology of the body or the sensorimotor contingencies (proprioceptive, visual and acoustic). This process is meant to be open-ended, progressing in continuous operation from initially simple representations towards abstract concepts and categories similar to those used by humans. Our work focuses on the study of various techniques for:

- autonomous multimodal dimensionality reduction and concept discovery;
- incremental discovery and learning of objects using vision and active exploration, as well as of auditory speech invariants;
- learning of dictionaries of motion primitives with combinatorial structures, in combination with linguistic description;
- active learning of visual descriptors useful for action (e.g. grasping).

**Embodiment and maturational constraints** FLOWERS studies how adequate morphologies and materials (i.e. morphological computation), associated to relevant dynamical motor primitives, can importantly simplify the acquisition of apparently very complex skills such as full-body dynamic walking in biped. FLOWERS also studies maturational constraints, which are mechanisms that allow for the progressive and controlled release of new degrees of freedoms in the sensorimotor space of robots.

**Discovering and abstracting the structure of sets of uninterpreted sensors and motors** FLOWERS studies mechanisms that allow a robot to infer structural information out of sets of sensorimotor channels whose semantics is unknown, for example the topology of the body and the sensorimotor contingencies (proprioceptive, visual and acoustic). This process is meant to be open-ended, progressing in continuous operation from initially simple representations to abstract concepts and categories similar to those used by humans.

**Emergence of social behavior in multi-agent populations** FLOWERS studies how populations of interacting learning agents can collectively acquire cooperative or competitive strategies in challenging simulated environments. This differs from "Social learning and guidance" presented above: instead of studying how a learning agent can benefit from the interaction with a skilled agent, we rather consider here how social behavior can spontaneously emerge from a population of interacting learning agents. We focus on studying and modeling the emergence of cooperation, communication and cultural innovation based on theories in behavioral ecology and language evolution, using recent advances in multi-agent reinforcement learning.

**Cognitive variability across Lifelong development and (re)educational Technologies** Over the past decade, the progress in the field of curiosity-driven learning generates a lot of hope, especially with regard to a major challenge, namely the inter-individual variability of developmental trajectories of learning, which is particularly critical during childhood and aging or in conditions of cognitive disorders. With the societal purpose of tackling of social inequalities, FLOWERS deals to move forward this new research avenue by exploring the changes of states of curiosity across lifespan and across neurodevelopmental conditions (neurotypical vs. learning disabilities) while designing new educational or rehabilitative technologies for curiosity-driven learning. The information gaps or learning progress, and their awareness are the core mechanisms of this part of research program due to high value as brain fuel by which the individual's internal intrinsic state of motivation is maintained and leads him/her to pursue his/her cognitive efforts for acquisitions /rehabilitations. Accordingly, a main challenge is to understand these mechanisms in order to draw up supports for the curiosity-driven learning, and then to embed them into (re)educational technologies. To this end, two-ways of investigations are carried out in real-life setting (school, home, work place etc): 1) the design of curiosity-driven interactive systems for learning and their effectiveness study ; and 2) the automated personnalization of learning programs through new algorithms maximizing learning progress in ITS.

## 4 Application domains

**Neuroscience, Developmental Psychology and Cognitive Sciences** The computational modelling of life-long learning and development mechanisms achieved in the team centrally targets to contribute to our understanding of the processes of sensorimotor, cognitive and social development in humans. In particular, it provides a methodological basis to analyze the dynamics of the interaction across learning and inference processes, embodiment and the social environment, allowing to formalize precise hypotheses and later on test them in experimental paradigms with animals and humans. A paradigmatic example of this activity is the Neurocuriosity project achieved in collaboration with the cognitive neuroscience lab of Jacqueline Gottlieb, where theoretical models of the mechanisms of information seeking, active learning and spontaneous exploration have been developed in coordination with experimental evidence and investigation, *see*. Another example is the study of the role of curiosity in learning in the elderly, with a view to assessing its positive value against the cognitive aging as a protective ingredient (i.e. Industrial project with Onepoint and CuriousTECH associate team with M. Fernandes from the Cognitive neuroscience Lab of the University of Waterloo).

**Personal and lifelong learning assistive agents** Many indicators show that the arrival of personal assistive agents in everyday life, ranging from digital assistants to robots, will be a major fact of the 21st century. These agents will range from purely entertainment or educative applications to social companions that many argue will be of crucial help in our society. Yet, to realize this vision, important obstacles need to be overcome: these agents will have to evolve in unpredictable environments and learn new skills in a lifelong manner while interacting with non-engineer humans, which is out of reach of current technology. In this context, the refoundation of intelligent systems that developmental AI is exploring opens potentially novel horizons to solve these problems. In particular, this application domain requires advances in artificial intelligence that go beyond the current state-of-the-art in fields like deep learning. Currently these techniques require tremendous amounts of data in order to function properly, and they are severely limited in terms of incremental and transfer learning. One of our goals is to drastically reduce the amount of data required in order for this very potent field to work when humans are in-the-loop. We try to achieve this by making neural networks aware of their knowledge, i.e. we introduce the concept of uncertainty, and use it as part of intrinsically motivated multitask learning architectures, and combined with techniques of learning by imitation.

**Educational technologies that foster curiosity-driven and personalized learning.** Optimal teaching and efficient teaching/learning environments can be applied to aid teaching in schools aiming both at increase the achievement levels and the reduce time needed. From a practical perspective, improved models could be saving millions of hours of students' time (and effort) in learning. These models should also predict the achievement levels of students in order to influence teaching practices. The challenges of the school of the 21st century, and in particular to produce conditions for active learning that are personalized to the student's motivations, are challenges shared with other applied fields. Special education for children with special needs, such as learning disabilities, has long recognized the difficulty of personalizing contents and pedagogies due to the great variability between and within medical conditions. More remotely, but not so much, cognitive rehabilitative carers are facing the same challenges where today they propose standardized cognitive training or rehabilitation programs but for which the benefits are modest (some individuals respond to the programs, others respond little or not at all), as they are highly subject to inter- and intra-individual variability. The curiosity-driven technologies for learning and STIs could be a promising avenue to address these issues that are common to (mainstream and specialized) education and cognitive rehabilitation.

**Automated discovery in science.** Machine learning algorithms integrating intrinsically-motivated goal exploration processes (IMGEPs) with flexible modular representation learning are very promising directions to help human scientists discover novel structures in complex dynamical systems, in fields ranging from biology to physics. The *automated discovery* project lead by the FLOWERS team aims to boost the efficiency of these algorithms for enabling scientist to better understand the space of dynamics of bio-physical systems, that could include systems related to the design of new materials or new drugs with applications ranging from regenerative medicine to unraveling the chemical origins of life. As an example, Grizou et al. [117] recently showed how IMGEPs can be used to automate chemistry experiments addressing fundamental questions related to the origins of life (how oil droplets may self-organize into protocellular structures), leading to new insights about oil droplet chemistry. Such methods can be

applied to a large range of complex systems in order to map the possible self-organized structures. The *automated discovery* project is intended to be interdisciplinary and to involve potentially non-expert end-users from a variety of domains. In this regard, we are currently collaborating with Poietis (a bio-printing company) and Bert Chan (an independent researcher in artificial life) to deploy our algorithms. To encourage the adoption of our algorithms by a wider community, we are also working on an interactive software which aims to provide tools to easily use the automated exploration algorithms (e.g. curiosity-driven) in various systems.

**Human-Robot Collaboration.** Robots play a vital role for industry and ensure the efficient and competitive production of a wide range of goods. They replace humans in many tasks which otherwise would be too difficult, too dangerous, or too expensive to perform. However, the new needs and desires of the society call for manufacturing system centered around personalized products and small series productions. Human-robot collaboration could widen the use of robot in this new situations if robots become cheaper, easier to program and safe to interact with. The most relevant systems for such applications would follow an expert worker and works with (some) autonomy, but being always under supervision of the human and acts based on its task models.

**Environment perception in intelligent vehicles.** When working in simulated traffic environments, elements of FLOWERS research can be applied to the autonomous acquisition of increasingly abstract representations of both traffic objects and traffic scenes. In particular, the object classes of vehicles and pedestrians are of interest when considering detection tasks in safety systems, as well as scene categories ("scene context") that have a strong impact on the occurrence of these object classes. As already indicated by several investigations in the field, results from present-day simulation technology can be transferred to the real world with little impact on performance. Therefore, applications of FLOWERS research that is suitably verified by real-world benchmarks has direct applicability in safety-system products for intelligent vehicles.

## 5 Social and environmental responsibility

### 5.1 Footprint of research activities

AI is a field of research that currently requires a lot of computational resources, which is a challenge as these resources have an environmental cost. In the team we try to address this challenge in two ways:

- by working on developmental machine learning approaches that model how humans manage to learn open-ended and diverse repertoires of skills under severe limits of time, energy and compute: for example, curiosity-driven learning algorithms can be used to guide agent's exploration of their environment so that they learn a world model in a sample efficient manner, i.e. by minimizing the number of runs and computations they need to perform in the environment;
- by monitoring the number of CPU and GPU hours required to carry out our experiments. For instance, our work [11] used a total of 2.5 cpu years. More globally, our work uses large scale computational resources, such as the Jean Zay supercomputer platform, in which we use several hundred thousands hours of GPU and CPU each year.

### 5.2 Impact of research results

Our research activities are organized along two fundamental research axis (models of human learning and algorithms for developmental machine learning) and one application research axis (involving multiple domains of application, see the Application Domains section). This entails different dimensions of potential societal impact:

- **Towards autonomous agents that can be shaped to human preferences and be explainable** We work on reinforcement learning architectures where autonomous agents interact with a social partner to explore a large set of possible interactions and learn to master them, using language as a key communication medium. As a result, our work contributes to facilitating human intervention in the learning process of agents (e.g. digital assistants, video games characters, robots), which we believe is a key step towards more explainable and safer autonomous agents.

- **Reproducibility of research:** By releasing the codes of our research papers, we believe that we help efforts in reproducible science and allow the wider community to build upon and extend our work in the future. In that spirit, we also provide clear explanations on the statistical testing methods when reporting the results.
- **Digital transformation and Competences' challenges facing schools in the 21st century.** We expect our findings to inform the broader societal challenges inherent to the School of the 21st Century, ranging from helping children (and their teachers) to develop cross-domain skills for learning such as curiosity and meta-cognition, while improving inclusivity in schools (learners with disabilities, especially cognitive disabilities) as well as promoting lifelong learning in older adults (successful aging), using cognitive-based research findings.
- **AI and personalized educational technologies to reduce inequalities due to neurodiversity** The Flowers team develops AI technologies aiming to personalize sequences of educational activities in digital educational apps: this entails the central challenge of designing systems which can have equitable impact over a diversity of students and reduce inequality in academic achievement. Using models of curiosity-driven learning to design AI algorithms for such personalization, we have been working to enable them to be positively and equitably impactful across several dimensions of diversity: for young learners or for aging populations; for learners with low initial levels as well as for learners with high initial levels; for "normally" developing children and for children with developmental disorders; and for learners of different socio-cultural backgrounds (e.g. we could show in the KidLearn project that the system is equally impactful along these various kinds of diversities).
- **Health: Bio-printing** The Flowers team is studying the use of curiosity-driven exploration algorithm in the domain of automated discovery, enabling scientists in physics/chemistry/biology to efficiently explore and build maps of the possible structures of various complex systems. One particular domain of application we are studying is bio-printing, where a challenge consists in exploring and understanding the space of morphogenetic structures self-organized by bio-printed cell populations. This could facilitate the design and bio-printing of personalized skins or organoids for people that need transplants, and thus could have major impact on the health of people needing such transplants.
- **Tools for human creativity and the arts** Curiosity-driven exploration algorithms could also in principle be used as tools to help human users in creative activities ranging from writing stories to painting or musical creation, which are domains we aim to consider in the future, and thus this constitutes another societal and cultural domain where our research could have impact.
- **Education to AI** As artificial intelligence takes a greater role in human society, it is of foremost importance to empower individuals with understanding of these technologies. For this purpose, the Flowers lab has been actively involved in educational and popularization activities, in particular by designing educational robotics kits that form a motivating and tangible context to understand basic concepts in AI: these include the Inirobot kit (used by >30k primary school students in France ([see](#))) and the Poppy Education kit ([see](#)) now supported by the Poppy Station educational consortium ([see](#)).
- **Health: optimization of intervention strategies during pandemic events** Modelling the dynamics of epidemics helps proposing control strategies based on pharmaceutical and non-pharmaceutical interventions (contact limitation, lock down, vaccination, etc). Hand-designing such strategies is not trivial because of the number of possible interventions and the difficulty to predict long-term effects. This task can be cast as an optimization problem where state-of-the-art machine learning algorithms such as deep reinforcement learning, might bring significant value. However, the specificity of each domain – epidemic modelling or solving optimization problem – requires strong collaborations between researchers from different fields of expertise. Due to its fundamental multi-objective nature, the problem of optimizing intervention strategies can benefit from the goal-conditioned reinforcement learning algorithms we develop at Flowers. In this context, we have developed EpidemioOptim, a Python toolbox that facilitates collaborations between researchers in epidemiology and optimization ([see](#)).



## 6 Highlights of the year

- **Understanding human curiosity and metacognition.** H. Sauzéon and PY. Oudeyer started a new European project in collaboration with cognitive neuroscience labs of M. Gruber's in Univ. Cardiff, and Y. Fandakova's in University of Trier, aiming to study the joint development of curiosity and metacognition in adolescents, through a set of behavioural and neuro-imaging studies. This project also aims to leverage new insights to be applied in educational technologies. We collaborated with G. Molinaro and A. Collins (Univ. Berkeley, who visited us for 6 months) in a project where we developed and tested a new model of learning-progress driven curiosity-driven exploration, based on latent evaluation of learning progress: this work was published in [Neurips 2024](#). We also collaborated with M. Fernandes' group in Univ. Waterloo for a study on the role of curiosity on spatial memory (published in [Frontiers in Cognition](#)). Collaborating with A. Tricot (Univ. Montpellier), we developed a theoretical perspective to study the links between intrinsic motivation and cognitive load in the context of extended-reality educational interventions (published in the [British Journal of Educational Technologies](#)). Finally, we collaborated with M. Derex (IAST, Toulouse) for designing protocols to study the role of curiosity in social learning and in cross-cultural settings, to be deployed in 2025.
- **Autotelic Generative AI for Self-Improving Program Synthesis** In the context of the LLM4Code Inria challenge, the team started collaboration on projects exploring projects at the intersection of machine learning, generative models and program synthesis. In particular, C. Moulin-Frier, L. Teodorescu, C. Colas, J. Pourcel, T. Boulet and PY. Oudeyer developed collaborations with N. Fijalkow (Labri, CNRS), X. Hinault (Mnemozyne), G. Baudart (PiCube). In a first milestone project, we developed a method (ACES) enabling large language models to self-generate diverse and challenging programming puzzles (using autotelic exploration). This enables to generate difficult benchmarks that adapt to evolving capacities of frontier models without the need of human manual creation of new benchmark problems. At the same time, these self created problems also allow generative models to self-improve, leveraging interaction with e.g. a Python interpreter. ACES was [published](#) at Neurips 2024.
- **Grounding agentic LLMs in interactive environments** We continued collaborating with O. Sigaud, L. Soulier, N. Tome (Sorbonne Université) and S. Lamprier (Univ. Angers) on developing techniques enabling to use, adapt and ground LLMs as agent policies in interactive environments. We further developed online RL algorithms in this context, adapting SAC to the GLAM architecture (presented at the [IMOL workshop](#) at Neurips 2024). This also led to a paper studying the role of prompt overfitting and how to mitigate it (published at [NAACL 2025](#)).
- **Collective intelligence and social learning in AI systems** We continued to actively develop this research axis, resulting in several key collaborations and publications. In the context of the PhD thesis of Grgur Kovač, we published two journal papers in collaboration with Peter-Ford Dominey (INSERM), Remy Portelas (Ubisoft) and Masataka Sawayama (Univ. Tokyo) on leveraging insights from developmental psychology towards artificial socio-cultural agents, resulting in two journal papers [\[35, 36\]](#). We also continued our collaboration with Eleni Nisioti and Sebastian Risi (ITU Copenhagen) and Ida Momennejad (Microsoft Research), now studying collective innovation in LLM agents, resulting in a paper at the ALife conference [\[52\]](#). Finally, in collaboration with Maxime Derex (IAST Toulouse), Cédric Colas (MIT) and Gaia Molinaro (UC Berkeley), we developed an experimental paradigm to study cultural evolution in chains of LLMs [\[67\]](#) (recently accepted at the ICLR conference).
- **Education, generative AI and cognitive training** We conducted a study in middle schools to study whether schoolchildren understand and know how to use generative AI tools in the context of educational exercises, showing strong limits and pointing to two needs: training their metacognition and their AI literacy. Also, we developed software library tools enabling to use LLMs to partially automate qualitative analysis methods in social sciences, leveraging our prior work in this direction [\[190\]](#), opening new perspectives for studying qualitatively large text corpuses or verbal data from psychology or educational experiments. We also continued working on

evaluating the use of adaptive personalization algorithms (in particular ZPDES, based on the learning progress theory) for cognitive training, and with diverse populations. This was associated to a review of AI-based approaches to cognitive training, published in [Plos One](#).

- **Curiosity-driven AI for assisted scientific discovery:** We continued studying how curiosity-driven AI algorithms can enable scientists (physicists, chemists, biologists, etc) explore and map the space of self-organized behaviours in diverse complex systems [114]. In particular, through a collaboration with M. Levin at Tufts University, we studied **how autotelic AI systems (IMGEP algorithms) can enable cost effective discovery of diverse sophisticated and robust behaviors in gene regulatory networks**, resulting in a milestone paper published in [eLife](#) ([33]). In another project, we continued our work using exploration algorithms to study **self-organized structures in continuous CAs like Lenia**, enabling to discover self-organization of forms of primitive agency.
- **Software** The team continued to develop several key software libraries: [Lamorel](#), enabling LLMs to be used as agents in interactive environments; [AdTool](#), enabling easy use of autotelic exploration algorithms for automated discoveries in physics/chemistry/Alife; [Vivarium](#), for building and running multi-agents simulations using Jax, with a focus on educational use.
- **Popular science book on generative AI** D. Roy and P-Y. Oudeyer wrote a popular science book to introduce generative AI (mechanisms, applications, societal dimensions) to adolescents, as well as to their teachers and families. It is entitled "C'est (pas) moi, c'est l'IA", and was published in september 2024 by Nathan. It was reviewed in widely distributed magazines (e.g. Magazine de l'APPEL) and radios (e.g. France Culture, RFI). The web page of the book is [here](#).
- **Educational video series on generative AI** A. Torres-Legueta, C. Romac, T. Carta and P-Y. Oudeyer produced the pedagogical video series "ChatGPT explained in 5 mn", aimed at training generative AI literacy in a wide diversity of students (e.g. high school), available here: [Web page](#). They are under a Creative Commons licence, CC-BY, enabling open and free reuse. They were already integrated in the MOOC AI4T ([Web page](#)), as well as in an internal training platform of "Académie du Numérique du Ministère de la défense", in a mobile app made by Inria with educational materials related to AI ([Web page](#)), and are being adapted and integrated in a training platform for the whole population of civil servants in France, coordinated by DINUM.
- **Popular science articles** C. Romac, T. Carta and P-Y. Oudeyer published a popular science article in Pour la Science article (French version of Scientific American) on how to ground generative AI models in interactive environments: Romac, C., Carta, T., Oudeyer, P-Y., "L'IA face au réel", Mars 2024. C. Moulin-Frier, G. Hamon and P-Y. Oudeyer published a popular science article in Pour la Science on using AI to explore the origins of evolutionary processes in Alife systems: Moulin-Frier, C., Hamon, G., Oudeyer, P-Y., "Quand l'IA explique les prémices d'une vie", August 2024.
- **Hackatons** C. Moulin-Frier co-organized the 3rd edition of [Hack10bo](#)
- **Support to public policy** H. Sauzéon and P-Y. Oudeyer were interviewed and wrote reports to contribute to the [report of French Senate on AI and education](#). P-Y. Oudeyer made a series of interventions to support public policy makers and educational institutions to address the societal and educational challenges and opportunities provided by AI, and in particular generative AI. In particular, he made presentations at [Cnesco](#), [DINUM](#) (to train senior officials), BPI, CCIC-Unesco.
- **Conference organizations** H. Sauzéon and C. Mazon co-organized the scientific days on "Handicap and autonomy", as well as the scientific days of the Fedhra. Cédric Colas, Laetitia Teodorescu and Gaia Molinaro co-organized the [Intrinsically Motivated Open-ended Learning workshop](#) at Neurips 2024. P-Y. Oudeyer co-organized the Life, Structure and Cognition symposium 2024 at IHES, France, this year on the topic "A Multitude of Times", [Web Page](#).

## 6.1 Awards

Didier Roy and Pierre-Yves Oudeyer obtained the ["special prize"](#) from Académie Nationale des Sciences, Belles Lettres et Arts de Bordeaux

Marion Pech obtained the prize from Foundation Clément Fayat (TOPAZ project) for AI-based digital therapy ([Web page](#))

## 7 New software, platforms, open data

### 7.1 New software

#### 7.1.1 SocialAI

**Name:** SocialAI: Benchmarking Socio-Cognitive Abilities in Deep Reinforcement Learning Agents

**Keywords:** Artificial intelligence, Deep learning, Reinforcement learning, Large Language Models

**Functional Description:** Source code for the paper <https://arxiv.org/abs/2107.00956>.

A suite of environments for testing socio-cognitive abilities of artificial agents. Environments can be used in the multimodal setting (suitable for RL agents) and in the pure text setting (suitable for Large Language Model-based agents). Also contains RL and LLM baselines.

**URL:** <https://gitlab.inria.fr/gkovac/act-and-speak>

**Contact:** Grgur Kovac

#### 7.1.2 AutoDisc

**Keyword:** Complex Systems

**Functional Description:** AutoDisc is a software built for automated scientific discoveries in complex systems (e.g. self-organizing systems). It can be used as a tool to experiment automated discovery of various systems using exploration algorithms (e.g. curiosity-driven). Our software is fully Open Source and allows user to add their own systems, exploration algorithms or visualization methods.

**URL:** <https://gitlab.inria.fr/cromac/AutomatedDiscoveryTool>

**Contact:** Clément Romac

#### 7.1.3 ADTool

**Keywords:** Machine learning, Python, Cellular automaton, Physical simulation, Pattern discovery, Exploration

**Functional Description:** ADTool is a versatile and open-source Python framework designed to explore complex parametric systems using IMGEP algorithms (Intrinsic Motivation for Goal Exploration Processes) as described in <https://arxiv.org/pdf/1708.02190>. This curiosity-driven approach enables automatic exploration and the discovery of new behaviors across a wide range of domains, offering a novel way to study complex systems.

With ADTool, users can explore cellular automata such as Lenia, Particle Lenia, and Flowlenia to uncover patterns and emergent behaviors. Its capabilities extend to drug discovery, exploring chemical spaces to identify promising protein-ligand affinity profiles. The framework also ventures into physics, with applications such as searching for trajectories in the N-body problem, simulating the Kuramoto model, exploring the Gray-Scott reaction-diffusion system, and studying hypergraph rewriting systems for Wolfram physics. In digital art, ADTool fosters creativity by exploring processes like subtractive sound synthesis and other artistic methods.

The framework is designed to be flexible and extensible, allowing users to define their own systems and integrate custom exploration strategies. It includes mechanisms for saving discoveries to disk, making it easier to resume experiments or share results with collaborators. Additionally, an integrated visualization tool provides a user-friendly interface to track exploration progress, enhancing the understanding and analysis of results.

The scientific foundation of ADTool lies in "curiosity-search" algorithms, which autonomously explore behavioral spaces to identify interesting phenomena without predefined objectives. These algorithms, initially developed for robotic learning, are now applied to the study of emergent behaviors in various systems.

Whether you are a physicist, chemist, biologist, or digital artist, ADTool can help you explore and understand complex systems.

Reproducibility is guaranteed with a predefined Python environment and experiments can be launched with a simple command line: `python3 run.py -config_file examples/grayscott/gray_scott.json`

**Contact:** Zacharie Bugaud

#### 7.1.4 Kids Ask

**Keywords:** Human Computer Interaction, Cognitive sciences

**Functional Description:** Kids Ask is a web-based educational platform that involves an interaction between a child and a conversational agent. The platform is designed to teach children how to generate curiosity-based questions and use them in their learning in order to gain new knowledge in an autonomous way.

**URL:** <https://github.com/RaniaAbdelghani/KidAsk>

**Contact:** Rania Abdelghani

#### 7.1.5 ToGather

**Keywords:** Education, Handicap, Environment perception

**Scientific Description:** With participatory design methods, we have designed an interactive website application for educational purposes. This application aims to provide interactive services with continuously updated content for the stakeholders of school inclusion of children with specific educational needs.

**Functional Description:** Website gathering information on middle school students with neurodevelopmental disorders. Authentication is required to access the site's content. Each user can only access the student file(s) of the young person(s) they are accompanying. A student file contains 6 tabs, in which each type of user can add, edit or delete information: 1. Profile: to quickly get to know the student 2. Skills: evaluation at a given moment and evolution over time 3. Compendium of tips: includes psycho-educational tips 4. Meetings: manager and reports 5. News: share information over time 6. Contacts: contact information for stakeholders The student only has the right to view information about him/her.

**Publication:** [hal-03898867](https://hal.archives-ouvertes.fr/hal-03898867)

**Contact:** Cécile Mazon

**Participants:** Isabeau Saint-Supery, Cécile Mazon, Eric Meyer, Hélène Sauzéon

#### 7.1.6 mc\_training

**Name:** Platform for metacognitive training

**Keywords:** Human Computer Interaction, Education

**Functional Description:** This is a web platform for children between 9 and 11 years old, designed to help children practice 4 metacognitive skills that are thought to be involved in curiosity-driven learning: - the ability to identify uncertainties - the ability to generate informed hypotheses - the ability to ask questions - the ability to evaluate the value of a preconceived inference.

Children work on a reading-comprehension tasks and, for each of these skills, the platform offers help through a "conversation" with conversational agents that give instructions to perform the task, with respect to every skill, and can give suggestions if the child asks for it.

**Contact:** Rania Abdelghani

### 7.1.7 Evolution of adaptation mechanisms in complex environments

**Name:** Plasticity and evolvability under environmental variability: the joint role of fitness-based selection and niche-limited competition

**Keywords:** Evolution, Ecology, Dynamic adaptation

**Functional Description:** This is the code accompanying our paper "Plasticity and evolvability under environmental variability: the joint role of fitness-based selection and niche-limited competition" which is to be presented at the Gecco 2022 conference.

In this work we have studied the evolution of a population of agents in a world where the fitness landscape changes with generations based on climate function and a latitudinal model that divides the world in different niches. We have implemented different selection mechanisms (fitness-based selection and niche-limited competition).

The world is divided into niches that correspond to different latitudes and whose state evolves based on a common climate function.

We model the plasticity of an individual using tolerance curves originally developed in ecology. Plasticity curves have the form of a Gaussian that capture the benefits and costs of plasticity when comparing a specialist (left) with a generalist (right) agent.

The repo contains the following main elements :

folder `source` contains the main functionality for running a simulation  
`scripts/run/reproduce_gecco.py` can be used to rerun all simulations in the paper  
`scripts/evaluate` contains scripts for reproducing figures. `reproduce_figures.py` will produce all figures (provided you have already run `scripts/run/reproduce_gecco.py` to generate the data)  
folder `projects` contains data generated from running a simulation  
How to run  
To install all package dependencies you can create a conda environment as:

```
conda env create -f environment.yml
```

All script executions need to be run from folder `source`. Once there, you can use `simulate.py`, the main interface of the codebase to run a simulation, For example:

```
python simulate.py --project test_stable --env_type stable --num_gens 300 --capacity 1000 --num_niches 10 --trials 10 --selection_type NF --climate_mean_init 2
```

will run a simulation with an environment with a climate function whose state is constantly 2 consisting of 100 niches for 300 generations and 10 independent trials. The maximum population size will be  $1000 \times 2$  and selection will be fitness-based (higher fitness means higher chances of reproduction) and niche limited (individuals reproduce independently in each niche and compete only within a niche),

You can also take a look at `scripts/run/reproduce_gecco.py` to see which flags were used for the simulations presented in the paper.

Running all simulations requires some days. You can instead download the data produced by running `scripts/run/reproduce_gecco.py` from this google folder and unzip them under the `projects` directory.

**URL:** <https://github.com/eleninisioti/ClimateAndLearning>

**Contact:** Eleni Nisioti

### 7.1.8 SAPIENS

**Name:** SAPIENS: Structuring multi-Agent toPology for Innovation through ExperieNce Sharing

**Keywords:** Reinforcement learning, Multi-agent

**Functional Description:** SAPIENS is a reinforcement learning algorithm where multiple off-policy agents solve the same task in parallel and exchange experiences on the go. The group is characterized by its topology, a graph that determines who communicates with whom.

All agents are DQNs and exchange experiences have the form of transitions from their replay buffers.

Using SAPIENS we can define groups of agents that are connected with others based on a a) fully-connected topology b) small-world topology c) ring topology or d) dynamic topology.

Install required packages You can install all required python packages by creating a new conda environment containing the packages in environment.yml:

```
conda env create -f environment.yml
```

And then activating the environment:

```
conda activate sapiens
```

Example usages Under notebooks there is a Jupyter notebook that will guide you through setting up simulations with a fully-connected and a dynamic social network structure for solving Wordcraft tasks. It also explains how you can access visualizations of the metrics produced during th\$

Reproducing the paper results Scripts under the scripts directory are useful for reproducing results and figures appearing in the paper.

With scripts/reproduce\_runs.py you can run all simulations presented in the paper from scratch.

This file is useful for looking at how the experiments were configured but better avoid running it: simulations will run locally and sequentially and will take months to complete.

Instead, you can access the data files output by simulations on this online repo.

Download this zip file and uncompress it under the projects directory. This should create a projects/paper\_done sub-directory.

You can now reproduce all visualization presented in the paper. Run:

```
python scripts/reproduce_visuals.py
```

This will save some general plots under visuals, while project-specific plots are saved under the corresponding project in projects/paper\_done

**URL:** <https://github.com/eleninisioti/SAPIENS>

**Contact:** Eleni Nisioti

### 7.1.9 architect-builder-abig

**Name:** Architect-Builder Iterated Guiding

**Keyword:** Artificial intelligence

**Functional Description:** Codebase for the paper Learning to guide and to be guided in the Architect-Builder Problem

ABIG stands for Architect-Builder Iterated Guiding and is an algorithmic solution to the Architect-Builder Problem. The algorithm leverages a learned model of the builder to guide it while the builder uses self-imitation learning to reinforce its guided behavior.

**URL:** <https://github.com/flowersteam/architect-builder-abig>

**Contact:** Tristan Karch

#### 7.1.10 EAGER

**Name:** Exploit question-Answering Grounding for effective Exploration in language-conditioned Reinforcement learning

**Keywords:** Reinforcement learning, Language, Question Generation Question Answering, Reward shaping

**Functional Description:** A novel QG/QA framework for RL called EAGER In EAGER, an agent reuses the initial language goal sentence to generate a set of questions (QG): each of these self-generated questions defines an auxiliary objective. Here, generating a question consists in masking a word of the initial language goal. Then the agent tries to answer these questions (guess the missing word) only by observing its trajectory so far. When it manages to answer a question correctly (QA) it obtains an intrinsic reward proportional to its confidence in the answer. The QA module is trained using a set of successful example trajectories. If the agent follows a path too different from correct ones at some point in its trajectory, the QA module will not answer the question correctly, resulting in zero intrinsic reward. The sum of all the intrinsic rewards measures the quality of a trajectory in relation to the given goal. In other words, maximizing this intrinsic reward incentivizes the agent to produce behaviour that unambiguously explains various aspects of the given goal.

**URL:** <https://github.com/flowersteam/EAGER>

**Contact:** Thomas Carta

#### 7.1.11 Flow-Lenia

**Name:** Flow Lenia: Mass conservation for the study of virtual creatures in continuous cellular automata

**Keywords:** Cellular automaton, Self-organization

**Functional Description:** This repo contains the code to run the Flow Lenia system which is a continuous parametrized cellular automaton with mass conservation. This work extends the classic Lenia system with mass conservation and allows to implement new feature like local parameter, environment components etc

Several declination of the system (1 or several channels etc ) are available

Please refer to the associated paper for the details of the system

Implemented in JAX

**URL:** <https://github.com/erwanplantec/FlowLenia>

**Contact:** Gautier Hamon

#### 7.1.12 Kidlearn: money game application

**Functional Description:** The games is instantiated in a browser environment where students are proposed exercises in the form of money/token games (see Figure 1). For an exercise type, one object is presented with a given tagged price and the learner has to choose which combination of bank notes, coins or abstract tokens need to be taken from the wallet to buy the object, with various constraints depending on exercises parameters. The games have been developed using web technologies, HTML5, javascript and Django.

**URL:** <https://flowers.inria.fr/research/kidlearn/>

**Contact:** Benjamin Clement



**Figure 1:** Four principal regions are defined in the graphical interface. The first is the wallet location where users can pick and drag the money items and drop them on the repository location to compose the correct price. The object and the price are present in the object location. Four different types of exercises exist: M : customer/one object, R : merchant/one object, MM : customer/two objects, RM : merchant/two objects.

### 7.1.13 cognitive-testbattery

**Name:** Cognitive test battery of human attention and memory

**Keywords:** Open Access, Cognitive sciences

**Scientific Description:** Cognitive test batteries are widely used in diverse research fields, such as cognitive training, cognitive disorder assessment, or brain mechanism understanding. Although they need flexibility according to the objectives of their usage, most of the test batteries are not available as open-source software and not be tuned by researchers in detail. The present study introduces an open-source cognitive test battery to assess attention and memory, using a javascript library, p5.js. Because of the ubiquitous nature of dynamic attention in our daily lives, it is crucial to have tools for its assessment or training. For that purpose, our test battery includes seven cognitive tasks (multiple-objects tracking, enumeration, go/no-go, load-induced blindness, task-switching, working memory, and memorability), common in cognitive science literature. By using the test battery, we conducted an online experiment to collect the benchmark data. Results conducted on two separate days showed the high cross-day reliability. Specifically, the task performance did not largely change with the different days. Besides, our test battery captures diverse individual differences and can evaluate them based on the cognitive factors extracted from latent factor analysis. Since we share our source code as open-source software, users can expand and manipulate experimental conditions flexibly. Our test battery is also flexible in terms of the experimental environment, i.e., it is possible to experiment either online or in a laboratory environment.

**Functional Description:** The evaluation battery consists of 6 cognitive activities (serious games: multi-object tracking, enumeration, go/no-go, Corsi, load-induced blindness, taskswitching, memorability). Easily deployable as a web application, it can be re-used and modified for new experiments. The tool is documented in order to facilitate the deployment and the analysis of results.

**URL:** <https://github.com/flowersteam/cognitive-testbattery>



**Publication:** [hal-03723887](#)

**Contact:** Maxime Adolphe

**Participants:** Pierre-Yves Oudeyer, H el ene Sauz eon, Masataka Sawayama, Maxime Adolphe

#### 7.1.14 Sensorimotor-lenia

**Keywords:** Cellular automaton, Gradient descent, Curriculum Learning

**Functional Description:** Source code for the search of sensorimotor agency in cellular automata associated to this blogpost <https://developmentalsystems.org/sensorimotor-lenia/>. The code allows to find rules in the cellular automata Lenia (through gradient descent, curriculum learning and diversity search) that lead to the self-organization of moving agents robust to perturbation by obstacles.

**URL:** <https://github.com/flowersteam/sensorimotor-lenia-search>

**Contact:** Gautier Hamon

#### 7.1.15 Lamorel

**Keywords:** Large Language Models, Reinforcement learning, Distributed computing

**Scientific Description:** Lamorel allows for seamless scaling of LLMs when using embodied artificial agents such as Reinforcement Learning agents. One can use and modify the LLM in any part of such agents (policy, goal sampler, social peer...). Lamorel is particularly useful when performing large-scale experiments on clusters.

It was already used in several papers, notably leading to the first paper performing online RL on an LLM-based agent in an embodied environment (Carta et. al, 2023).

**Functional Description:** Lamorel was initially designed to easily use LLMs in interactive environments. It is especially made for high throughput using a distributed architecture. The philosophy of \*Lamorel\* is to be very permissive and allow as much as possible usage of LLMs while maintaining scaling: the application should run with 1 or N LLMs.

For this reason, it is not specialised neither in RL nor in particular in RLHF. Our examples illustrate how \*Lamorel\* can be used for various applications including RLHF-like finetuning. However, one must understand that \*Lamorel\*'s philosophy means that users must implement themselves what they want to do with the LLM(s).

This is why we advise users knowing in advance they want to do RLHF, especially without any modification of classic implementations, to use libs specialised in RLHF that already come with RL implementations (e.g. RL4LMs, TRL). On the other hand, users more inclined to experiment with implementations or looking for an LLM lib they can use in different projects may prefer Lamorel.

Here are Lamorel's key features: 1. Abstracts the use of LLMs (e.g. tonekization, batches) into simple calls

2. Provides a method to compute the log probability of token sequences (e.g. action commands) given a prompt 3. Is made for scaling up your experiments by deploying multiple instances of the LLM and dispatching the computation thanks to a simple configuration file 4. Provides access to open-sourced LLMs from the Hugging Face's hub along with Model Parallelism to use multiple GPUs for an LLM instance 5. Allows one to give their own PyTorch modules to compute custom operations (e.g. to add new heads on top of the LLM) 6. Allows one to train the LLM (or part of it) thanks to a Data Parallelism setup where the user provides its own update method

**URL:** <https://github.com/flowersteam/lamorel>

**Publications:** [hal-03970122](#), [hal-04844089](#), [hal-04844077](#)

**Contact:** Cl ement Romac

### 7.1.16 GLAM

**Name:** Grounding LAnguage Models

**Keywords:** Large Language Models, Reinforcement learning

**Scientific Description:** Recent works successfully leveraged Large Language Models' (LLM) abilities to capture abstract knowledge about world's physics to solve decision-making problems. Yet, the alignment between LLMs' knowledge and the environment can be wrong and limit functional competence due to lack of grounding. In this paper, we study an approach (named GLAM) to achieve this alignment through functional grounding: we consider an agent using an LLM as a policy that is progressively updated as the agent interacts with the environment, leveraging online Reinforcement Learning to improve its performance to solve goals. Using an interactive textual environment designed to study higher-level forms of functional grounding, and a set of spatial and navigation tasks, we study several scientific questions: 1) Can LLMs boost sample efficiency for online learning of various RL tasks? 2) How can it boost different forms of generalization? 3) What is the impact of online learning? We study these questions by functionally grounding several variants (size, architecture) of FLAN-T5.

**Functional Description:** GLAM is a new approach to achieve alignment between a Large Language Model (LLM) and a considered environment/world through functional grounding: we consider an agent using an LLM as a policy that is progressively updated as the agent interacts with the environment, leveraging online Reinforcement Learning to improve its performance to solve goals.

**URL:** [https://github.com/flowersteam/Grounding\\_LLMs\\_with\\_online\\_RL](https://github.com/flowersteam/Grounding_LLMs_with_online_RL)

**Publication:** hal-03970122

**Contact:** Clément Romac

### 7.1.17 SBMLtoODEjax

**Keywords:** SBML, JAX, Python, Numerical simulations, Numerical optimization, Automatic differentiation, Ordinary differential equations, Biomedical data

**Scientific Description:** Advances in bioengineering and biomedicine demand a deep understanding of the dynamic behavior of biological systems, ranging from protein pathways to complex cellular processes. Biological networks like gene regulatory networks and protein pathways are key drivers of embryogenesis and physiological processes. Comprehending their diverse behaviors is essential for tackling diseases, including cancer, as well as for engineering novel biological constructs. Despite the availability of extensive mathematical models represented in Systems Biology Markup Language (SBML), researchers face significant challenges in exploring the full spectrum of behaviors and optimizing interventions to efficiently shape those behaviors. Existing tools designed for simulation of biological network models are not tailored to facilitate interventions on network dynamics nor to facilitate automated discovery. Leveraging recent developments in machine learning (ML), this paper introduces SBMLtoODEjax, a lightweight library designed to seamlessly integrate SBML models with ML-supported pipelines, powered by JAX. SBMLtoODEjax facilitates the reuse and customization of SBML-based models, harnessing JAX's capabilities for efficient parallel simulations and optimization, with the aim to accelerate research in biological network analysis.

**Functional Description:** SBMLtoODEjax extends SBMLtoODEpy, a python library developed in 2019 for converting SBML files into python files written in Numpy/Scipy. The chosen conventions for the generated variables and modules are slightly different from the standard SBML conventions (used in the SBMLtoODEpy library) with the aim here to accommodate for more flexible manipulations while preserving JAX-like functional programming style.

**URL:** <https://developmentalsystems.org/sbmltoodejax/>

**Publication:** [hal-04317246](#)

**Contact:** Mayalen Etcheverry

**Partner:** Tufts University

#### 7.1.18 Vivarium

**Name:** Large-scale simulator for research and teaching in Artificial Intelligence and Artificial Life

**Keywords:** Simulation, Artificial intelligence, Artificial Life, Multi-Agents System, Teaching of programming, Research

**Functional Description:** This project aims to seize these opportunities through the design and implementation of a software platform providing an integrated simulation environment for research, teaching, and dissemination in the fields of Artificial Intelligence (AI) and Artificial Life (AL). The project is titled The Vivarium, which reflects a fundamental aspect of the convergence between these two domains: the emergence of complex behaviors, whether in the natural or artificial world, necessarily relies on a need for adaptation to a complex environment in which many autonomous entities interact.

It will be used as an educational software in a course from CISC Master at UPF-Barcelona in January 2025.

**Release Contributions:** This release corresponds to the state of the repo after all fixes were made following the SDIC course at Universitat Pompeu Fabra of Barcelone (CSIM Master) in January 2025 .

This version mostly focuses on educational purposes, with ready-to-use practical sessions in notebooks/sessions. Corentin Léger was the main contributor over the last year.

**News of the Year:** Corentin Léger, ingénieur de recherche recruté sur l'ANR JCJC ECOCURL (porté par Clément Moulin-Frier) a mené un gros travail de développement du logiciel au cours de l'année 2024. Ses applications pour l'enseignement sont maintenant validées. Le logiciel a notamment été utilisé pendant 10 heures de travaux pratiques dans le Master CSIC de Universitat Pompeu Fabra à Barcelone, Espagne.

**URL:** <https://github.com/flowersteam/vivarium>

**Contact:** Clément Moulin-Frier

**Participants:** Corentin Leger, Clément Moulin-Frier, Martial Marzloff

#### 7.1.19 LLM\_Culture

**Keywords:** LLM, Multi-Agents System, Natural language processing

**Functional Description:** Code for the 'Cultural evolution in populations of Large Language Models' paper. This repository provides a comprehensive framework for studying the cultural evolution of linguistic content in populations of Large Language Models (LLM).

It allows organizing LLM agents into networks wherein each agent interacts with neighboring agents by exchanging texts. Each agent can be assigned specific personalities and transmission instructions, serving as prompts for generating new texts from their neighbors' narratives. Once the network structure and agent characteristics are defined, you can simulate the cultural evolution of texts across generations of agents. We also provide built-in metrics and visualizations to analyze the results.

**URL:** <https://github.com/flowersteam/LLM-Culture>

**Contact:** Jeremy Perez

**Participants:** Jeremy Perez, Corentin Leger

### 7.1.20 TelephoneGameLLMs

**Keywords:** Large Language Models, Multi-Agents System, Cultural Evolution

**Functional Description:** Code for the paper "When LLMs Play the Telephone Game: Cumulative Changes and Attractors in Iterated Cultural Transmissions" <https://arxiv.org/abs/2407.04503> In this paper, we introduce conceptual and methodological tools for evaluating Large Language Models in multi-turn settings. Those tools are inspired by cultural evolutionary theory, and in particular by the concepts of cultural attractors.

**URL:** <https://github.com/flowersteam/TelephoneGameLLM>

**Publication:** hal-04714994

**Contact:** Jeremy Perez

### 7.1.21 styr

**Name:** Stick To Your Role

**Keywords:** LLM, Cognitive sciences

**Functional Description:** Code for our paper <https://arxiv.org/abs/2402.14846> and leaderboard <https://huggingface.co/spaces/flowers-team/StickToYourRoleLeaderboard>.

Enables evaluating LLMs using personal value questionnaires (PVQ, SVS). More precisely, it instructs the LLM to simulated various personas and exposes it to different contexts (e.g. long reddit posts). Then it evaluates the value stability of the simulated population between those contexts. Additionally, it computes confirmatory factor analysis (CFI, SRMR, RMSEA), and the structure of expressed values (stress metric).

**URL:** [https://github.com/flowersteam/value\\_stability](https://github.com/flowersteam/value_stability)

**Contact:** Grgur Kovac

### 7.1.22 transformerXL\_PPO\_JAX

**Keywords:** Reinforcement learning, Transformer

**Functional Description:** This repository provides a JAX implementation of TransformerXL with PPO in a RL setup following : "Stabilizing Transformers for Reinforcement Learning" from Parisotto et al. (<https://arxiv.org/abs/1910.06764>).

The code uses the [PureJaxRL](#) template for PPO and copied some of the code from [hugging face transformer XL repo](#) transferring it to JAX. We also took inspiration from the pytorch code in <https://github.com/MarcoMeter/episodic-transformer-memory-ppo>, which has some simplification of gradient propagation and positional encoding compared to transformerXL as it is described in the original paper (<https://arxiv.org/abs/1901.02860>). The training handles [Gymnax](<https://github.com/RobertTLange/gymnax>) environment.

We also tested it on [Craftax](#), on which it beat the baseline presented in the paper (<https://arxiv.org/abs/2402.16801>) including PPO-RNN, training with unsupervised environment design and intrinsic motivation. Notably we reach the 3rd level (the sewer) and obtain several advanced advancements, which was not achieved by the methods presented in the paper. See [Craftax Results](#) for more informations.

The training of a 5M transformer on craftax for 1e9 steps (with 1024 environments) takes about 6h30 on a single A100.

**Contact:** Gautier Hamon

### 7.1.23 ER-MRL

**Keywords:** Reinforcement learning, Evolutionary Algorithms, Recurrent network

**Functional Description:** Code for the "Evolving-Reservoirs-for-Meta-Reinforcement-Learning" (ER-MRL) paper (<https://arxiv.org/abs/2312.06695>).

We adopt a computational framework based on meta reinforcement learning, modeling the interplay between evolution and development. At the evolutionary scale, we evolve reservoirs, a family of recurrent neural networks generated from hyperparameters. These evolved reservoirs are then utilized to facilitate the learning of a behavioral policy through reinforcement learning. This is done by encoding the environment state through the reservoir before providing it to the agent's policy.

**Contact:** Corentin Leger

## 7.2 New platforms

### 7.2.1 ToGather application

**Participants:** Cécile Mazon, Hélène Sauzéon, Eric Meyer, Isabeau Saint-Supery.

**Name:** Application for Specialized education

**Keywords:** Parent-professional relationships; user-centered design; school inclusion; autism spectrum disorder; ecosystemic approach

**Participants:** Isabeau Saint-supery, Cécile Mazon, Hélène Sauzéon, Agilonaute

**Scientific Description:** With participatory design methods, we have designed an interactive website application for educational purposes. This application aims to provide interactive services with continuously updated content for the stakeholders of school inclusion of children with specific educational needs. Especially, the services provide: 1) the student's profile with strengths and weaknesses; 2) an evaluation and monitoring over time of the student's repertoire of acquired, emerging or targeted skills; 3) a shared notebook of effective psycho-educational solutions for the student ; 4) a shared messaging system for exchanging "news" about the student and his/her family and, 5) a meeting manager allowing updates of evaluations (student progress). This application is currently assessed with a field study. Then, it will be transferred to the Academy of Nouvelle-Aquitaine-Bordeaux of the National Education Ministry.

**URL:** The website is not online yet, but all informations such as tutorials are [here](#).

**Publication:** [hal-03436355](#)

## 7.3 Open data

As part of the "Jack of All Trades, Master of Some, a Multi-Purpose Transformer Agent" project jointly done with Hugging Face and collaborators from INSA Lyon, we release the JAT dataset, the first dataset for generalist agent training. The JAT dataset contains hundreds of thousands of expert trajectories collected with expert agents.

## 8 New results

The team's research program, within the domain of developmental artificial intelligence, aims to study mechanisms of open-ended learning, and in particular the role of curiosity-driven autotelic learning and

the role of language as a cognitive tool. We study these topics both in humans and AI systems, both at the level of individuals and at the level of cultural groups, and both at the fundamental and application levels.

Here, we present our recent results along the following research dimensions:

- Open-ended learning and autotelic AI with large language models;
- Models of cultural evolution in humans and AI systems;
- An Eco-Evo-Devo perspective on Artificial Intelligence;
- Generative AI and educational technologies;
- Theories of human curiosity-driven learning
- Curiosity-driven educational technologies;
- Curiosity-driven AI for assisted scientific discovery;

## 8.1 Open-ended learning and autotelic AI with large language models

The team continued to lay the foundations of autotelic AI [105, 104], i.e. the science studying mechanisms enabling artificial agents to learn to represent and sample their own goals and achieve open-ended learning.

### 8.1.1 ACES: Generating a Diversity of Challenging Programming Puzzles with Autotelic Generative Models

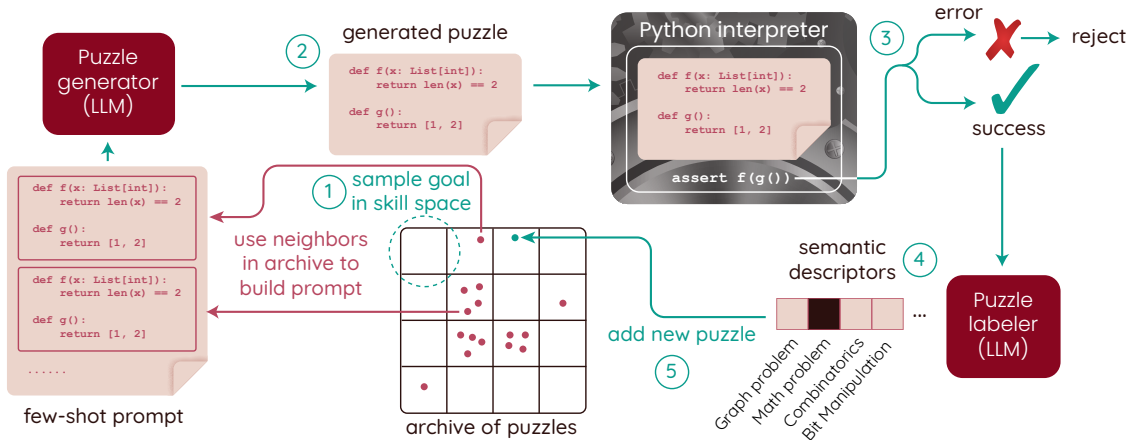
**Participants:** Julien Pourcel (*correspondant*), Cédric Colas, Gaia Molinaro, Pierre-Yves Oudeyer, Laetitia Teodorescu.

**Motivation.** In this project, we examine how one can generate an interesting diversity of programming puzzles (same domain as Codeplay). We recall that this is an important case study for linguistic autotelic agents because it is a first step towards generalist agents inventing their own problems. Inspired by the Evolution Through Large Models (ELM) method where authors evolve robot morphologies expressed as Sodarace programs using a Large Language Model as a mutation operator, we aim to develop an evolutionary method to create a diverse population of problems using pretrained Language Models. We remark that diversity-producing methods (such as Map-Elites) need a Behavioral Characterization (BC) space in which to measure the diversity of their evolved populations; this is feasible with virtual creatures but seems pretty hard with programming puzzles. We thus introduce the notion of a *Semantic* BC space, composed of abstract categories, and labelling inside this space is done through LLM responses. In our case, we introduce 10 programming descriptors:

- 0 - Sorting and Searching
- 1 - Counting and Combinatorics
- 2 - Trees and Graphs
- 3 - Mathematical Foundations
- 4 - Bit Manipulation
- 5 - String Manipulation
- 6 - Geometry and Grid Problems
- 7 - Recursion and Dynamic Programming
- 8 - Stacks and Queues
- 9 - Optimization Algorithms

We then define an archive of generated programming puzzles and their solutions, and the position of a puzzle in the archive is given by the combination of descriptors that the puzzle-solution pair belongs to (the semantic representation of a puzzle thus being a 10-dimensional vector). The semantic archive is used to store puzzles.

We then perform experiments with the following algorithms:



**Figure 2:** Overview of ACES. ACES maintains an archive of discovered puzzles grouped into cells indexed by their semantic representation (skill combination). ACES runs in several steps: 1) sample a target semantic goal and relevant examples from the archive. 2) given these, generate a puzzle  $f$  and its solution  $g$  with the puzzle generator. 3) test the validity of that pair by running `assert f(g())` in the interpreter. 4) if the pair is valid, obtain its semantic representation with the puzzle labeler. 5) add the new pair to its corresponding cell in the archive.

- **ACES:** our proposed method samples a target cell (combination of descriptors) in the archive at random and populates a few-shot prompt for the language model with puzzles from neighboring cells in the archive. See Figure 2 for an illustration.
- **ELM Semantic:** based on ELM, example puzzles and solutions are given as few-shot in-context examples and a puzzle sampled from the archive is then mutated.
- **ELM:** same as the previous one, except we do not use the semantic archive for sampling: instead we build an archive with centroidal voronoi tessellations, from the embedding of puzzles inside the latent space of a Language Model. This baseline allows us to compare the semantic archive with a more classical one;
- **Static Gen:** In this method, puzzles are sampled from the train set and added as few-shot examples in the prompt;

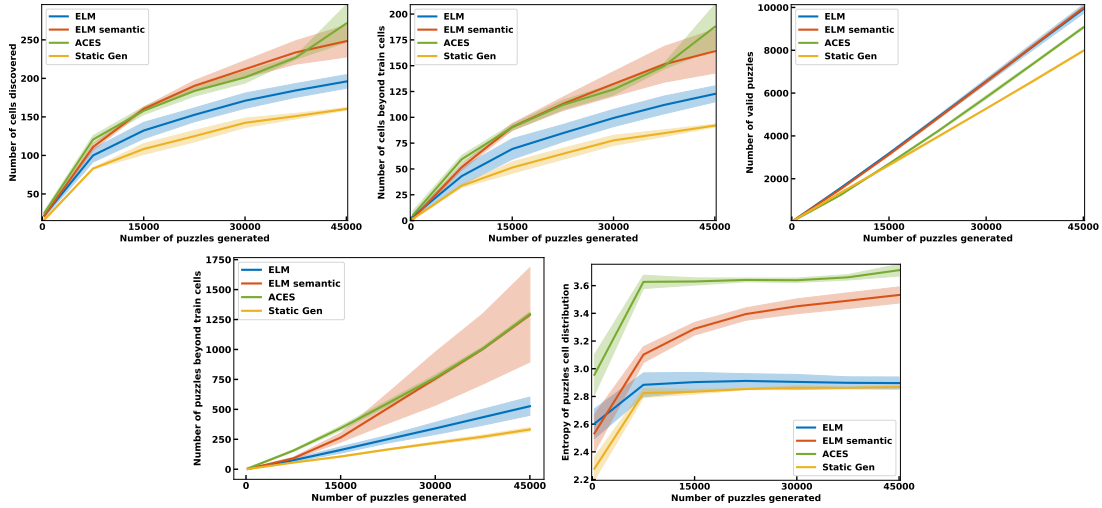
For all experiments we seed the archive with the P3 train set.

**Results.** We report results of our runs in Figure 3. Overall, the methods based on semantic archives, ACES and ELM-Semantic, achieve the highest diversity in the semantic space. We report diversity measures inside the embedding spaces of various smaller language models in Figure 4. In these figures we see that overall ACES outperforms other methods in this measure of diversity. We additionally perform tests of the suitability of generated puzzles as finetuning data for smaller LMs. For all methods, we finetune a smaller model (OpenLlama-3b) on the generated set and we test the  $\text{pass}@k$  metric for different values of  $k$  on the P3 test set; we report the scores in Figure 5. From that figure we see that we encounter a tradeoff between how diverse the data is and how useful it is to get a high score on the P3 test set. Further work is needed to get data that is both diverse and useful.

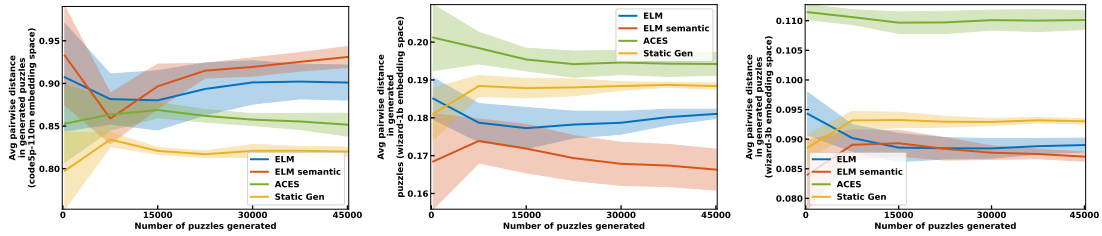
### 8.1.2 MAGELLAN: Metacognitive Generalization of Learning Progress for Online RL in LLM agents

**Participants:** Loris Gaven (*correspondant*), Thomas Carta, Clément Romac, Cédric Colas, Pierre-Yves Oudeyer, Olivier Sigaud (*ISIR Sorbonne Université, Paris, France*), Sylvain Lamprier (*Univ Angers, LERIA*).

We are developing **MAGELLAN**, a method designed to enable LLM-based reinforcement learning (RL) agents to estimate their own Learning Progress (LP) and use it to dynamically organize their training



**Figure 3:** Diversity of generated puzzles in semantic space. We report the evolution of several diversity metrics computed in the semantic space as a function of the number of puzzle-solution pairs generated by the puzzle generator. Semantic algorithms (algame and ELM semantic) achieve higher diversity in the semantic space.



**Figure 4:** Diversity of generated puzzles in embedding spaces. We report the evolution of the pairwise distance between puzzle-solution pair embeddings as a function of the number of generated puzzle-solution pairs, for three different embedding representation spaces (average across seeds).

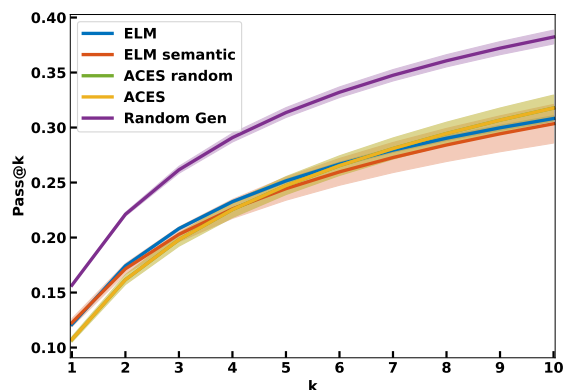
curriculum. By leveraging the LLM’s rich semantic representations, MAGELLAN allows agents to generalize LP estimations to unseen, language-defined goals, overcoming limitations of classical methods that require direct evaluation of each goal.

MAGELLAN uses the LLM to generate latent representations of goals and tasks, capturing their semantic relationships. It continuously monitors the agent’s performance over time, estimating LP as the change in success rates for specific goals. This approach enables the agent to identify goals where it is making progress and focus its training on those areas. MAGELLAN’s integration ensures that the LLM-based agent can simultaneously refine its policy and competence estimations, adapting both to new tasks in real time.

Our experiments in the **Little-Zoo environment**, which features hierarchical and commonsense-driven tasks, demonstrate that MAGELLAN effectively prioritizes high-LP goals, even when faced with novel or unseen tasks. Unlike traditional LP estimation methods, which rely on direct evaluations and struggle with generalization, MAGELLAN enables the agent to quickly identify meaningful learning opportunities. This results in faster adaptation, improved sample efficiency, and more effective curriculum organization, paving the way for truly autonomous agents capable of navigating vast and complex goal spaces.

### 8.1.3 SAC-GLAM: Improving Online RL for LLM agents with Soft Actor-Critic and Hindsight Relabeling





**Figure 5:** Downstream performance on the P3 test set. Pass@k is the fraction of puzzles solved after  $k$  attempts ( $k \in [1:10]$ ). Green overlaps with yellow.

**Participants:** Loris Gaven (*correspondant*), Thomas Carta, Clément Romac, Pierre-Yves Oudeyer, Olivier Sigaud (*ISIR Sorbonne Université, Paris, France*), Sylvain Lamprier (*Univ Angers, LERIA*).

We developed **SAC-GLAM**, a novel reinforcement learning approach designed for Large Language Model (LLM) agents. By adapting Soft Actor-Critic (SAC) with Hindsight Experience Replay (HER), SAC-GLAM addresses key limitations of on-policy methods like PPO-GLAM, enabling more efficient and effective learning in complex environments.

SAC-GLAM leverages HER to reframe failed attempts as learning opportunities by relabeling goals based on achieved outcomes, significantly improving sample efficiency. Additionally, the method integrates pre-trained LLM policies with a randomly initialized critic, incorporating a warm-up phase and architectural enhancements to ensure stable and effective training. We also used  $n$ -step returns to further enhance sample efficiency.

Our approach was evaluated in **Playground-Text**, a text-based environment featuring sequential and multi-goal tasks. Results showed that SAC-GLAM outperformed PPO-GLAM in sample efficiency while maintaining comparable time performance, even when extended with HER.

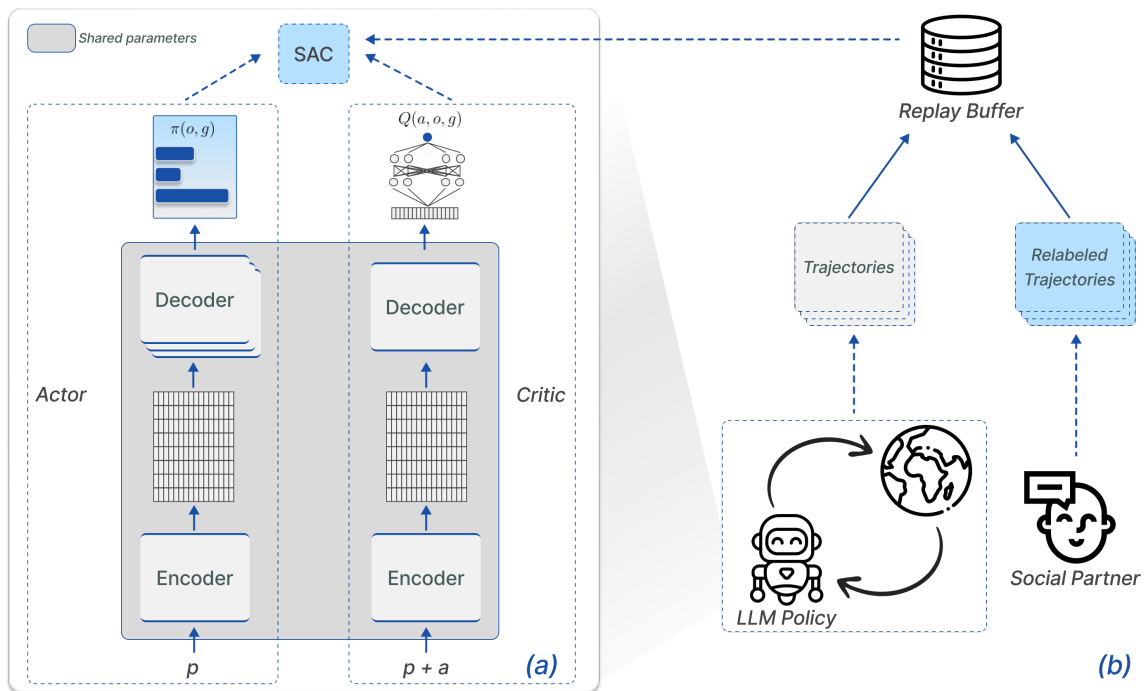
This work represents a step forward in creating autonomous, self-motivated LLM agents capable of setting and pursuing their own goals, leveraging language understanding to learn more effectively. It highlights the potential of combining off-policy RL techniques with LLMs to tackle complex, multi-goal tasks in text-based environments.

#### 8.1.4 Learning when (not) to use tools with Memory-based In-context learning and online multi-objective RL in LLM agents

**Participants:** Clément Romac (*correspondant*), Pierre-Yves Oudeyer.

Augmenting Large Language Models (LLMs) with tools they can call is now seen as a natural way to enhance their abilities. In most prior work, very large LLMs were used and calling tools was not assumed costly in the light of the model’s size. In this paper, we argue that using tools is particularly important for smaller-scale LLMs, and that it should be assumed as a costly decision regarding the lower compute needs from such small models. We propose to study, in such conditions, the central question of *when* should tools be used, which we show depends on 1) a model’s ability to solve tasks without tools and 2) how costly tools are.

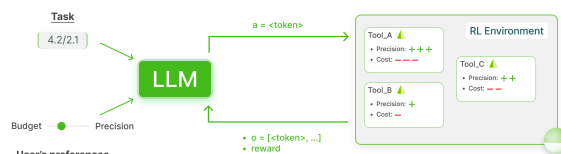
We take the perspective of an autonomous agent, where the LLM should explore and discover by itself when tools are helpful for it to minimize its answers’ errors and tool usage (see Figure 7). We show that



**Figure 6: The SAC-GLAM method.** (a) depicts the agent’s architecture when an encoder-decoder LLM is used: the actor computes an action probability as the probability computed by the LLM of the action’s tokens to follow the observation and goal concatenated in the prompt  $p$ , while the critic computes the Q-value for each action  $a$  and the prompt  $p$  with an MLP attached to the decoder’s last hidden state. (b) illustrates the agent-environment interaction, where trajectories are generated and added to the replay buffer. We used an environment where a social partner relabels these trajectories with hindsight goals.

current methods fall short of finding the optimal strategy when tools are costly. Notably, we argue this comes both from an inability to properly explore (as identified by prior work) and an inability to reinforce good strategies quickly.

We introduce a new method that aims to solve these two issues. We take inspiration from Episodic Control and use an external memory to store and retrieve the best trajectories. When appended to the prompt, we show these trajectories, along with the in-context learning abilities of LLMs, allowing the model to both try new strategies and reproduce the best ones depending on the instruction. However, Episodic Control alone easily gets stuck in local optima using deceptive strategy. To overcome this, we use an off-policy online RL method named AWR. While necessitating as few changes on the LLM as current methods used (e.g. PPO), AWR’s off-policy nature allows the agent to reinforce best strategies and escape from local optima efficiently.



**Figure 7:** We study how LLMs can autonomously learn how and when to use tools when solving tasks.

### 8.1.5 WorldLLM: Improving LLMs’ world modelling using Curiosity-driven Theory-making

**Participants:** Guillaume Levy, Clément Romac (*correspondant*), Thomas Carta, Cédric Colas, Pierre-Yves Oudeyer.

Large Language Models (LLMs) possess various knowledge about the world that allows them to answer questions or propose plans for solving tasks. Yet, this knowledge can be wrong or incomplete, especially when considering problems with specific properties (e.g. counterintuitive). To this day, aligning the world modelling abilities of LLMs remains a central and open challenge. Prior work notably showed that allowing LLMs to perform embodied experiences can improve their world model [96].

Learning a world model of an environment has been extensively studied in the Reinforcement Learning (RL) community. A large body of work focused on the forward model aspect, that is, predicting what will happen (i.e. the next observation) given an action and what happened before (i.e. current observation). While learning such a forward model can allow better planning or policy improvement without interacting with the environment, it has also been shown useful to foster exploration. Notably, ICM [161] showed how the prediction error of a learned forward model can be used as an intrinsic reward motivating the policy to explore and find transitions unknown by the world model.

Parallel to this, Theory-based RL [185] proposes to integrate human-inspired inductive reasoning to produce a world model able to predict causal relations based on intuitive theories. The use of intuitive theories notably allows humans to generalize their world model and quickly adapt it by improving their theories.

In this paper, we propose to study how to improve the world modelling abilities of an LLM. Inspired by Theory-based RL and following recent work, we propose to see the LLM as our causal model of the environment (i.e. the forward model) and to search for natural language theories that would improve its ability to predict the environment’s dynamics. We argue that generating rules that help predict the world fosters abstract and general rules that humans can interpret. To obtain such rules, we rely on the inductive reasoning abilities of LLMs together with Bayesian Inference. Inspired by ICM, we also propose to actively collect the experiences used to induce rules by leveraging an RL agent solely trained on an intrinsic reward provided by the LLM-based forward model’s error.

To summarize, we introduce **WorldLLM**, a method for improving an LLM’s world modelling capabilities. It is constituted of 1) a forward model leveraging the LLM we want to improve, 2) another LLM incrementally inducing theories using a Bayesian Inference framework, and 3) an RL agent that

looks for experiences the LLM-based forward model fails to predict given the current theories. Inspired by how humans and, in particular, scientists explain the world, our method iteratively proposes theories, designs experiments to challenge them and refines the theories according to the experiments' outcome. We show that, upon perfect data collectors replacing our RL agent, our framework generates rules that significantly improve an LLM's ability to predict the outcome of actions. But more importantly, we show that random data collection is insufficient to cover the space of possible outcomes. By inciting the RL agent to find transitions the forward model does not predict well, we show our LLM successfully model the space of possible outcomes.

### 8.1.6 Reinforcement Learning for Aligning Large Language Models Agents with Interactive Environments: Quantifying and Mitigating Prompt Overfitting

**Participants:** Mohamed Salim Aissi (*ISIR Sorbonne Université, Paris, France*), Clément Romac (*correspondant*), Thomas Carta, Sylvain Lamprier (*Univ Angers, LERIA*), Pierre-Yves Oudeyer, Olivier Sigaud (*ISIR Sorbonne Université, Paris, France*), Laure Soulier (*ISIR Sorbonne Université, Paris, France*), Nicholas Thome (*ISIR Sorbonne Université, Paris, France*).

In GLAM [96], we showed an environment's dynamics could be explicitly integrated into an LLM using Reinforcement Learning (RL) to fine-tune the LLM to solve sequential decision-making tasks. While this showed one could ground an LLM's abilities at manipulating words related to environments' dynamics, it remains unclear how RL impacts the inherent knowledge of the LLM and whether the model gains generalization capabilities.

In this paper, we study the sensitivity of LLMs to prompt formulations and its impact on knowledge acquisition. We define a set of different prompt formulations to evaluate LLM performance when prompts are changed, and also analyze how the LLM represents these prompts. Our findings reveal that the performance of LLMs is highly sensitive to prompt variations, suggesting that fine-tuning only induces superficial updates and fails to improve the acquisition of new knowledge about the environment. By analogy with observational overfitting in RL, we refer to this phenomenon as *prompt overfitting*. With this in mind, our paper proposes two main contributions:

- We design experiments to measure prompt overfitting issues of LLMs in interactive environments. The study reveals that fine-tuned LLMs heavily depend on the training prompts, exhibiting a significant drop in "zero-shot" performance when using new prompt formulations. To further analyze this behavior, we thoroughly analyze latent representations and salient tokens in LLMs, both showing a strong bias towards the prompt formulation.
- We design experiments to measure prompt overfitting issues of LLMs in interactive environments. The study reveals that fine-tuned LLMs heavily depend on the training prompts, exhibiting a significant drop in "zero-shot" performance when using new prompt formulations. To further analyze this behavior, we thoroughly analyze latent representations and salient tokens in LLMs, both showing a strong bias towards the prompt formulation.
- We propose a solution for mitigating prompt overfitting with a contrastive regularization loss that makes the latent representation of the LLM invariant to prompt formulations. This solution significantly improves the zero-shot performance and the robustness to prompt variations, as well as the acquisition of new knowledge about the environment. See Figure 8.

### 8.1.7 Jack of All Trades, Master of Some, a Multi-Purpose Transformer Agent

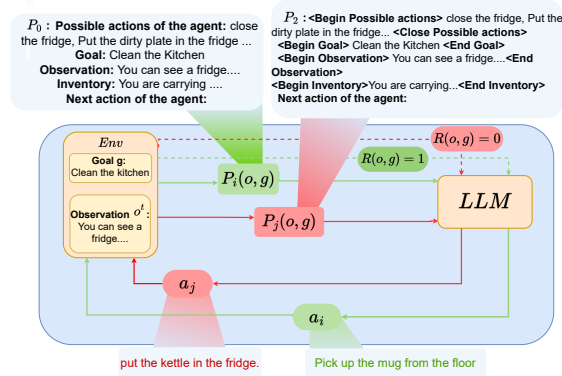


Figure 8: Our constrastive RL framework.

**Participants:** Clément Romac (*correspondant*), Quentin Gallouédec (*Ecole Centrale de Lyon, CNRS, Université Claude Bernard Lyon 1, INSA Lyon, Université Lumière Lyon 2, LIRIS, UMR5205, 69130 Ecully, France*), Edward Beeching (*Hugging Face*), Emmanuel Dellandréa (*Ecole Centrale de Lyon, CNRS, Université Claude Bernard Lyon 1, INSA Lyon, Université Lumière Lyon 2, LIRIS, UMR5205, 69130 Ecully, France*).

We present the Jack of All Trades (JAT) model, a transformer-based generalist agent capable of high performance on tasks ranging from Atari to MuJoCo, while being able to generate text and image captions. Inspired by GATO [165], we collect what is, to our knowledge, the largest open dataset for training generalist agents. This dataset features hundreds of thousands of expert trajectories.

We then trained a transformer model with Behavioral Cloning on these trajectories. To overcome the challenge of unifying the different observation and action spaces of our various environments, we designed a new approach, simpler than previous methods, using a single embedding space and one token per action or observation (see Figure 9). We also propose to add an auxiliary loss turning our JAT model into a forward model (i.e. predicting the next observation given the previous observation and action). We show this auxiliary task improves our model’s abilities. Finally, we also train our model using a Language Modeling loss on text generation and image captioning datasets.

We show our model achieves strong performance on all the tasks it was trained with Behavioral Cloning. These results are above human-level performance on multiple tasks and on par with results reported by GATO. Our JAT model also exhibits rudimentary language modelling abilities. These results can be explained by the small-scale datasets used for text generation and image captioning.

We open-source our dataset along with the experts used to collect trajectories as well as the trained JAT model. We believe our work can act as a foundation model for future agents while opening up new avenues in designing generalist agents.

### 8.1.8 LLM-based goal generation for autotellic exploration with goal-conditioned RL

**Participants:** Guillaume Pourcel (*correspondant*), Grgur Kovač, Thomas Carta, Cédric Colas, Pierre-Yves Oudeyer.

Autotellic agents, capable of autonomously generating and pursuing their own goals, represent a promising approach to open-ended learning and skill acquisition in reinforcement learning. Such agents learn to set and pursue their own goals. This challenge is even more difficult in open worlds that require inventing new previously unobserved goals. In this work, we propose an architecture where a single generalist autotellic agent is trained on an automatic curriculum of goals. We leverage a large language models (LLMs) to generate goals as code for reward functions based on learnability and

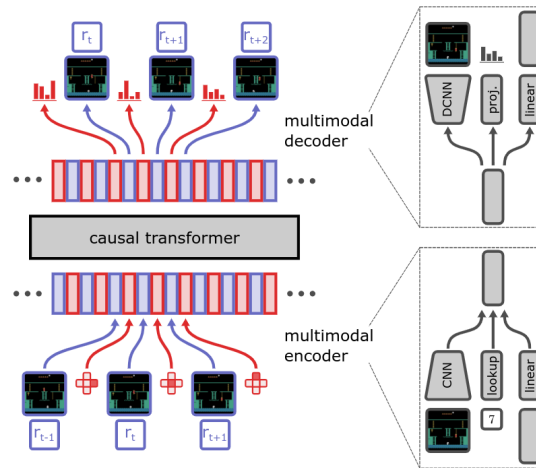


Figure 9: Our JAT model's architecture.

difficulty estimates. The goal-conditioned RL agent is trained on those goals sampled based on learning progress. We evaluate our method in terms of novelty, difficulty, and learnability of the generated goals. Our preliminary experiments imply that our method generates goals of increasing difficulty while maintaining their learnability and diversity.

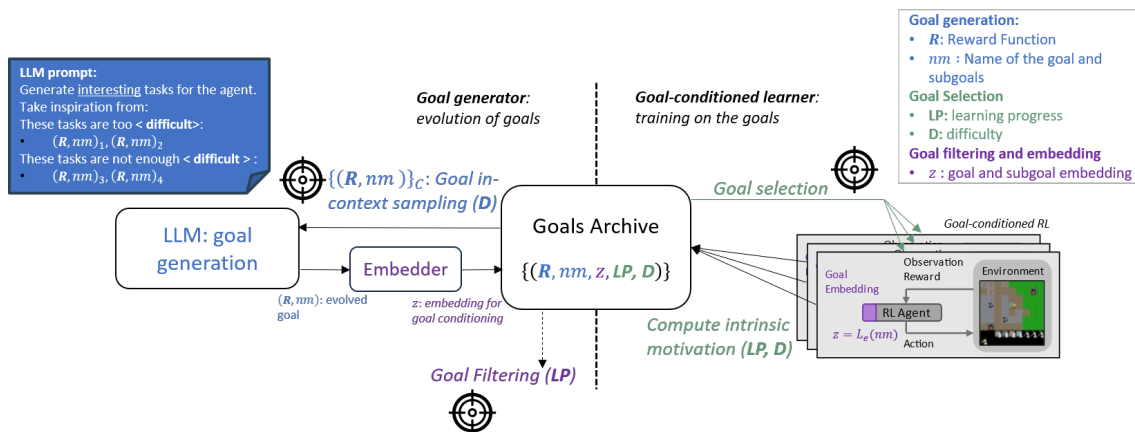


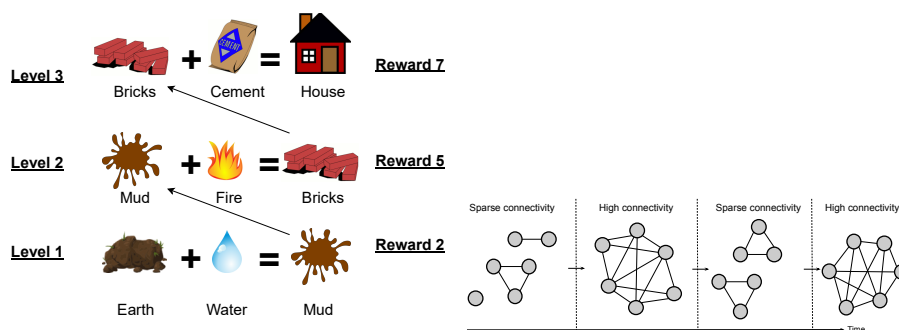
Figure 10: The used architecture.

## 8.2 Models of cultural evolution in humans and AI systems

As **generative AI systems become powerful cultural transmission technologies** that influence human cultural evolution in important ways, and can also have their own cultural processes through machine-machine large scale interaction, the study of the dynamics of cultural processes in populations of AI systems/humans becomes crucial.

### 8.2.1 The effect of social network structure on collective innovation

**Participants:** Eleni Nisioti (*correspondant*), Mateo Mahaut, Pierre-Yves Oudeyer, Ida Momennejad, Sebastian Risi, Pierre-Yves Oudeyer, Clément Moulin-Frier.



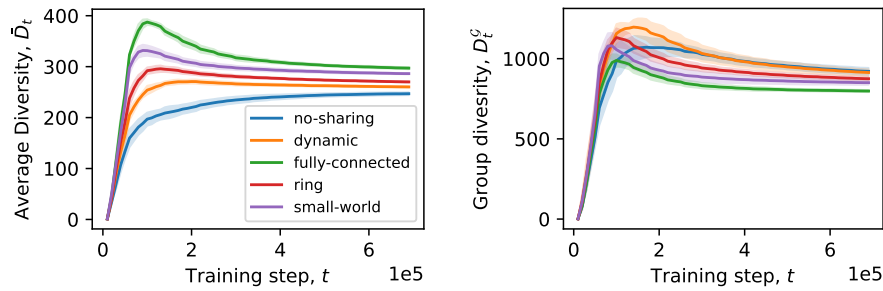
**Figure 11:** (Left) Illustration of an innovation task, consisting of an initial set of elements (Earth, Water) and a recipe book indicating which combinations create new elements. Upon creating a new element the player moves up an innovation level and receives a reward that increases monotonically with levels. (Right) Dynamic social network structures oscillate between phases of low connectivity, where experience sharing takes place within clusters, and high connectivity, where experiences spread between clusters.

Innovations are a central component of open-ended skill acquisition: they denote the emergence of new solutions by the recombination of existing ones and their presence is necessary to ensure a continuous complexification of an agent’s cultural repertoire. While we often tend to attribute discoveries to certain innovative individuals, if we shed a broad perspective at the history of our species we see that human innovation is primarily a collective process. Fields such as psychology and anthropology have been studying the ability of human groups to innovate for some time, with studies indicating that the social network structure has a significant impact: fully-connected structures are better suited for quick convergence in easy problems with clear global optima, while partially-connected structures perform best in difficult tasks where local optima may lure agents away from the globally optimal solution [111]. At the same time a parallel story is unfolding in reinforcement learning (RL): distributed RL is a sub-field where multiple agents solve a task collectively [151]. Compared to the single-agent paradigm, distributed RL algorithms converge quicker and often achieve superior performance. However, these algorithms have only considered full connectivity. In this inter-disciplinary project, we presented a novel learning framework that augments distributed RL with the notion of a social network structure and employed it to study the hypothesis from human studies that partial connectivity performs best in innovation tasks.

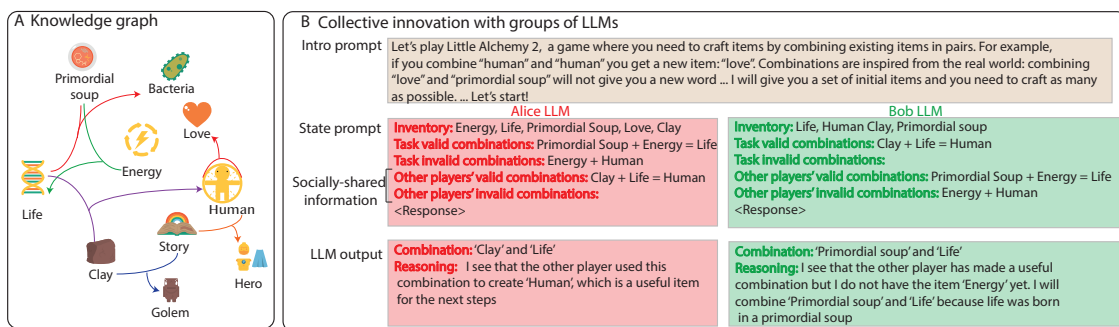
**Cultural evolution in populations of RL agents.** We implemented such innovation tasks using *Wordcraft*, a recently introduced RL playground inspired from the *Little Alchemy 2 game* (see left of figure 11 for an illustration of how this task works). We considered a wide diversity of social network structures: static structures that remain constant throughout learning (fully-connected, ring, small-world) and a dynamic structure where the group oscillates between phases of low and high connectivity (we illustrate this dynamic structure on the right of figure 11). Each agent in our implementation employs the DQN learning algorithm and exchanges experiences that have the form of sequences of state-action combinations with its neighbors.

A central conclusion of our empirical analysis was that the dynamic social network structure performs best. In addition to the performance groups achieve we measured behavioral and mnemonic metrics such as behavioral conformity and mnemonic diversity. Such metrics were inspired from human studies and helped us further analyze the behavior of groups. For example, one empirical observation was that sharing experiences did not help the group learn quicker in a very simple innovation task; instead the fully-connected group was the slowest. By looking at the diversity in the memories of the agents we observed that the fully-connected structure had the highest individual diversity (left of figure 12) and the lowest group diversity (right of figure 12): sharing experiences with others diversifies an individual’s experiences but also homogenizes the group, which is bad for its performance.

We see the contribution of this project as two-fold. From the perspective of fields studying human intelligence, we have shown that using RL algorithms as computational tool is a promising direction towards increasing the verisimilitude of simulations and analyzing both behavior and memory. From the perspective of RL, we have shown that distributed RL algorithm should move beyond the fully-connected



**Figure 12:** (Left) Illustration of an innovation task, consisting of an initial set of elements (Earth, Water) and a recipe book indicating which combinations create new elements. Upon creating a new element the player moves up an innovation level and receives a reward that increases monotonically with levels. (Right) Dynamic social network structures oscillate between phases of low connectivity, where experience sharing takes place within clusters, and high connectivity, where experiences spread between clusters.



**Figure 13:** Studying collective innovation in groups of LLMs: A) we experiment with Little Alchemy 2, a game where players combine real-world items to create new ones. A knowledge graph describes the possible combinations (we only present a small sub-part of the graph which contains 720 items in total) B) Alice-LLM and Bob-LLM are two LLMs playing the game together. They are provided with the same intro prompt, explaining the rules of the game, and the same task (they start with the same set of items). Alice-LLM and Bob-LLM have identical weights but behave differently because the state prompt depends on their crafting history. They are informed about the actions of others through their prompt. In this paper, we study how groups of such LLM agents are able to efficiently explore a knowledge graph, focusing in particular on the effect of different social structures specifying with whom and when they can share information

architecture and explore groups with dynamic topologies. This work is currently a preprint [154] and is about to be submitted in PNAS. We open-source the code [at this link](#).

**Cultural evolution in populations of LLM agents.** In 2024, we have extended this framework with agents equipped with Large Language Models (LLMs) playing Little Alchemy 2, a creative video game originally developed for humans (figure 13). We, first, study an LLM in isolation and discover that it exhibits both useful skills and crucial limitations. We, then, study groups of LLMs that share information related to their behaviour and focus on the effect of social connectivity on collective performance. In agreement with previous human and computational studies (including the one described above), we observe that groups with dynamic connectivity out-compete fully-connected groups. Our work reveals opportunities and challenges for future studies of collective innovation that are becoming increasingly relevant as Generative Artificial Intelligence algorithms and humans innovate alongside each other. We published this work at the ALife 2024 conference [52].

## 8.2.2 The SocialAI School: Insights from Developmental Psychology Towards Artificial Socio-Cultural Agents



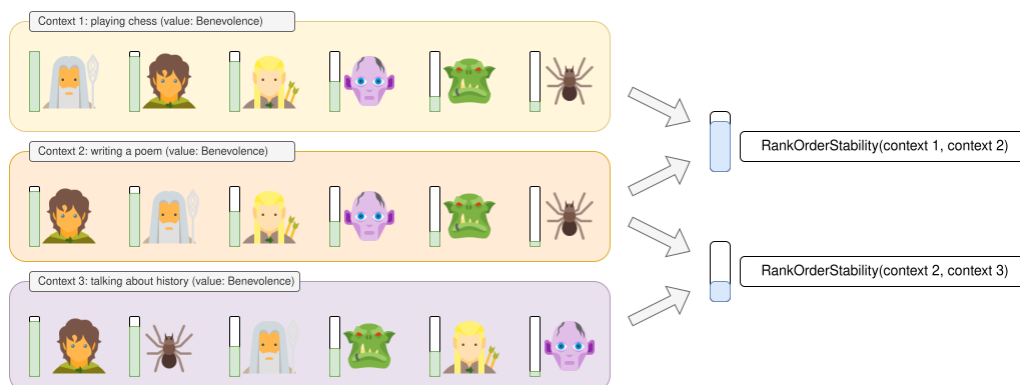
**Participants:** Grgur Kovač (*correspondant*), Remy Portelas, Peter Ford Dominey, Pierre-Yves Oudeyer.

This project aims to present a tool facilitating research into socio-cognitive AI agents. The project was finished in 2023 and published in 2024.

### 8.2.3 Stick to your Role! Stability of Personal Values Expressed in Large Language Models

**Participants:** Grgur Kovač (*correspondant*), Jeremy Perez, Remy Portelas, Masataka Sawayama, Peter Ford Dominey, Pierre-Yves Oudeyer.

Following our perspective paper [133] from last year we continued this project with an empirical paper [47], and then with a Leaderboard[73]. The standard way to study Large Language Models (LLMs) with benchmarks or psychology questionnaires is to provide many different queries from similar minimal contexts (e.g. multiple choice questions). However, due to LLMs' highly context-dependent nature, conclusions from such minimal-context evaluations may be little informative about the model's behavior in deployment (where it will be exposed to many new contexts). We argue that context-dependence (specifically, value stability) should be studied as a specific property of LLMs and used as another dimension of LLM comparison (alongside others such as cognitive abilities, knowledge, or model size). In the empirical paper [47], we present a case-study on the stability of value expression over different contexts (simulated conversations on different topics) as measured using a standard psychology questionnaire (PVQ) and on behavioral downstream tasks. Reusing methods from psychology, we study Rank-order stability on the population (interpersonal) level, and Ipsative stability on the individual (intrapersonal) level. We consider two settings (with and without instructing LLMs to simulate particular personas), two simulated populations, and three downstream tasks. We observe consistent trends in the stability of models and model families - Mixtral, Mistral, GPT-3.5 and Qwen families are more stable than LLaMa-2 and Phi. The consistency of these trends implies that some models exhibit higher value stability than others, and that stability can be estimated with the set of introduced methodological tools. When instructed to simulate particular personas, LLMs exhibit low Rank-order stability, which further diminishes with conversation length. This highlights the need for future research on LLMs that coherently simulate different personas. This paper provides a foundational step in that direction, and, to our knowledge, it is the first study of value stability in LLMs. For the leaderboard[73], we extend the methodology with Confirmatory Factor Analysis and Stress metrics, which test for the structure of expressed values. Many new models were compared in the leaderboard, and it is regularly updated.



**Figure 14:** Rank-order stability - correlation of participants in different contexts (ordered in terms of expressing a certain value)

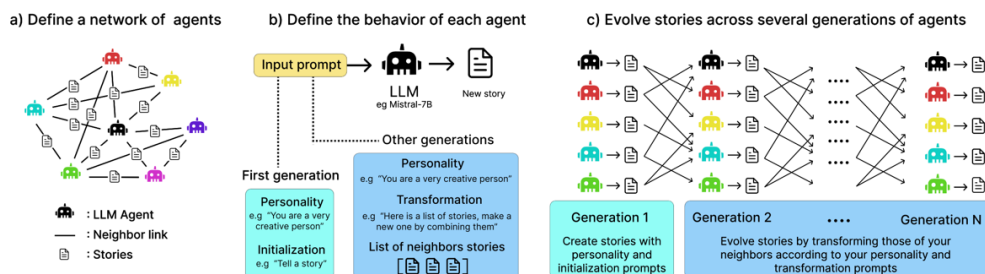
### 8.2.4 Cultural evolution in populations of Large Language Models

**Participants:** J r my Perez (*correspondant*), Corentin L ger, Marcela Ovando-Tellez, Chris Foulon, Joan Dussauld, Pierre-Yves Oudeyer, Cl ment Moulin-Frier.

Research in cultural evolution aims at providing causal explanations for the change of culture over time. Over the past decades, this field has generated an important body of knowledge, using experimental, historical, and computational methods. While computational models have been very successful at generating testable hypotheses about the effects of several factors, such as population structure or transmission biases, some phenomena have so far been more complex to capture using agent-based and formal models. This is in particular the case for the effect of the transformations of social information induced by evolved cognitive mechanisms. We here proposed that leveraging the capacity of Large Language Models (LLMs) to mimic human behavior may be fruitful to address this gap. On top of being an useful approximation of human cultural dynamics, multi-agents models featuring generative agents are also important to study for their own sake. Indeed, as artificial agents are bound to participate more and more to the evolution of culture, it is crucial to better understand the dynamics of machine-generated cultural evolution.

In this project, we presented a framework for simulating cultural evolution in populations of LLMs, allowing the manipulation of variables known to be important in cultural evolution, such as network structure, personality, and the way social information is aggregated and transformed. The software we developed for conducting these simulations is open-source and features an intuitive user-interface.

The model simulates the cultural evolution of linguistic content in a population LLMs. Each agent can be seen as an independent instance of a LLM. The agents are arranged according to a specified social network structure. At the first generation, all agents are prompted with an Initialization Prompt (e.g. ‘‘Tell me a story’’). All agents then output an answer by passing the initialization prompt to their respective instance of the LLM. The agents then transmit stories to their neighbors (according to the specified network structure): each agent receives a new prompt, which is the concatenation of a Transformation Prompt (e.g. ‘‘Combine the stories you received’’) and the list of stories produced by its neighbors at the previous generation. Agents can additionally be provided with a personality (e.g. ‘‘You are very imaginative’’).

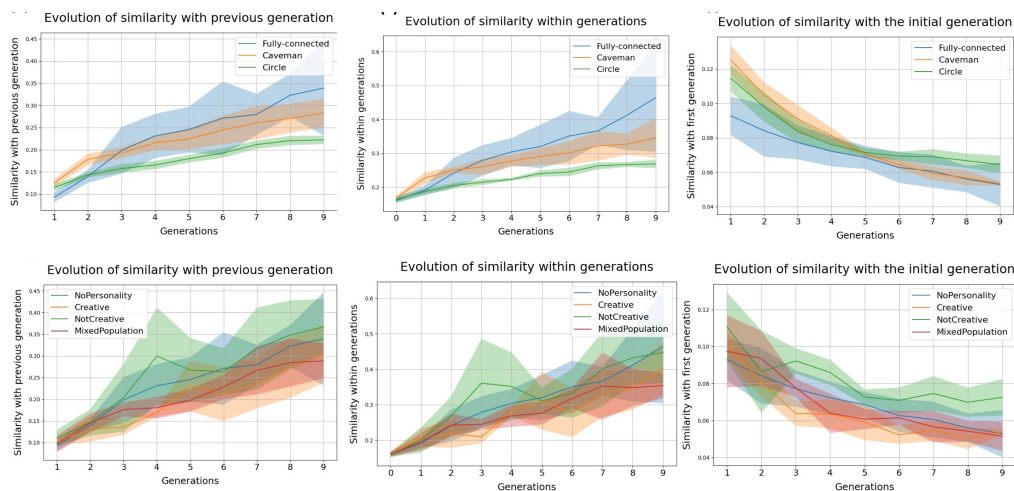


**Figure 15:** (a) LLM agents are organized into networks wherein each agent interacts with neighboring agents by exchanging stories. (b) Each agent is assigned a specific personality and either initialization instructions (for the first generation) or transformation instructions (after the first generation), serving as prompts for generating new stories from their neighbors’ narratives. (c) Once the network structure and agent characteristics are defined, we simulate the cultural evolution of texts across generations of agents. The simulation begins by prompting agents to initialize stories, after which we allow the narratives to evolve dynamically through interactions within the agent network

We also introduce several visualizations and measures that are useful for tracking the results of the simulations, namely the similarity matrix, the similarity graph, and the evolution of the words used at each generation.

We find some interesting dynamics, such as the fact that stories seem to evolve in a punctuated manner: there is an alternation of phases where stories are transmitted with no modifications, and phases where stories are modified. These dynamics recall previous observations from experiments and

modelling work in cultural evolution [24] [46] [35] [34]. The effect of network structure also mirrors results found in humans, with more efficient networks (that is, networks with smaller average path length) leading to quicker diffusion of information and thus lower diversity. Finally, we found that manipulating personality appears to affect cultural dynamics; for instance, assigning a “creative” personality to agents led groups to generate more diversity.



**Figure 16:** Effect of network structure (top row) and personality (bottom row) on cultural dynamics

Overall, this contribution shows how using generative agents to simulate cultural evolution may be promising, both for generating hypotheses about human cultural evolution and for better understanding the dynamics of machine-generated culture.

The article is available as a pre-print [68], and we open-source the code [here](#).

### 8.2.5 When LLMs Play the Telephone Game: Cumulative Changes and Attractors in Iterated Cultural Transmissions

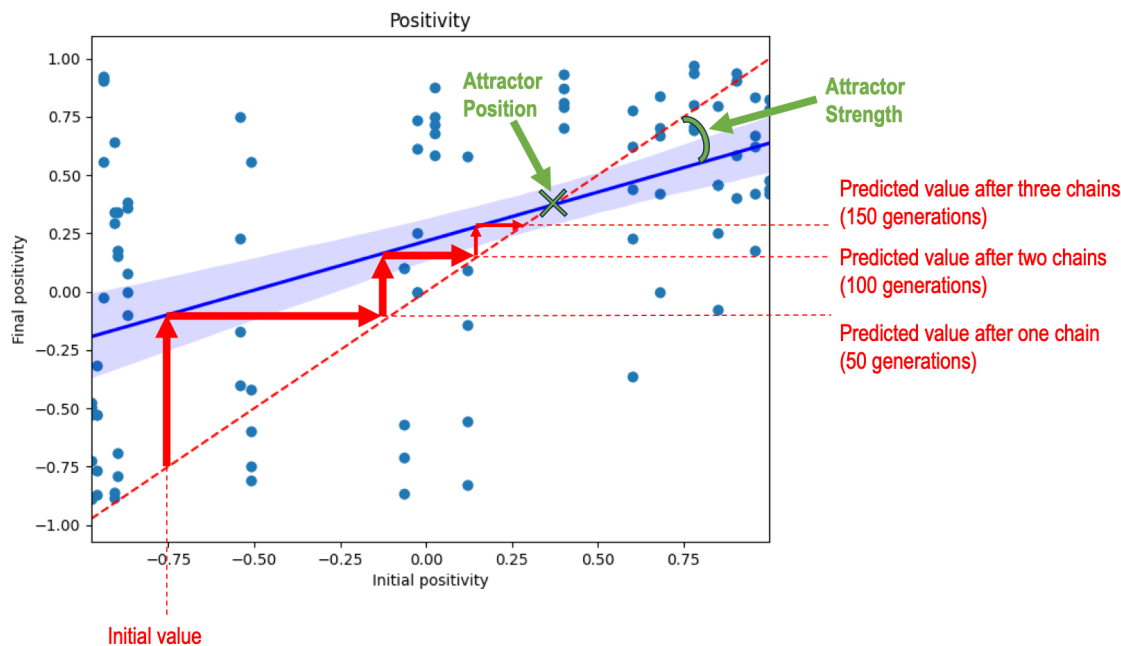
**Participants:** Jérémy Perez (*correspondant*), Grgur Kovač, Corentin Léger, Cédric Colas, Gaia Molinaro, Maxime Derex, Pierre-Yves Oudeyer, Clément Moulin-Frier.

As large language models (LLMs) start interacting with each other and generating an increasing amount of text online, it becomes crucial to better understand how information is transformed as it passes from one LLM to the next. While significant research has examined individual LLM behaviors, existing studies have largely overlooked the collective behaviors and information distortions arising from iterated LLM interactions. Small biases, negligible at the single output level, risk being amplified in iterated interactions, potentially leading the content to evolve towards attractor states.

In this project, we ran a series of telephone game experiments, applying a transmission chain design borrowed from the human cultural evolution literature: LLM agents iteratively receive, produce, and transmit texts from the previous to the next agent in the chain.

Our main contributions are:

- We propose that there might be a gap in current LLM evaluations methods (single-turn evaluations might not be suited to assess the properties of multi-turn interactions)
- We empirically confirm this hypothesis by showing that multi-turn interactions indeed often lead to distributions of text properties that are significantly different from what is observed after a single interaction.
- We introduce novel conceptual and methodological tools to fill this gap, grounded in research in cultural evolution, and in particular the concept of cultural attractor.



**Figure 17:** Method for estimating attractor strength and position.

- We showcase the potential of this method by applying it to compare the effect of different tasks, of different models, of temperature, and of fine-tuning on the properties of multi-turn interactions.
- We find several robust effects, such as the fact that less constrained tasks lead to stronger attractors, that some properties possess stronger attractors than others, and that fine-tuning can shift the position and modify the strength of attractors.

These findings highlight the importance of accounting for multi-step transmission dynamics and represent a first step towards a more comprehensive understanding of LLM cultural dynamics.

This work was presented during a 15-minutes talk given at the Cultural Evolution Society conference. The article is available as a pre-print [67], and we open-source the code [here](#). We also created a [website](#) featuring a Data Explorer tool, allowing to directly inspect the texts generated during our experiments. An extended version of the paper has been accepted at the ICLR'2025 conference (one of the three main international conferences in machine learning).

### 8.2.6 Collapse of diversity and quality in the cultural evolution of Large Language models

**Participants:** Grgur Kovač (*correspondant*), Jérémy Perez (*correspondant*), Remy Portelas, Peter Ford Dominey, Pierre-Yves Oudeyer.

LLMs are increasingly contributing to the creation of content on Internet. This creates a feedback loop as next generations of models will be trained on this generated, synthetic data. This phenomenon is receiving increasing interest, in particular because previous studies have shown that it may lead to drastic losses in performance and quality, an outcome commonly referred to as “model collapse” [175]. While studying model collapse and mitigation strategies is under active investigation, existing studies have so far ignored how features of the data that the model is trained to learn impact collapse. Indeed, different online domains contain data that possess different properties. Some domains may be associated with a comparatively higher ratio of synthetic-to-human data (e.g. GitHub), while others may contain data of comparatively lower quality (e.g. Reddit). If future specialized models are trained on some specific internet domains, we may therefore expect different rates and forms of model collapse depending on the

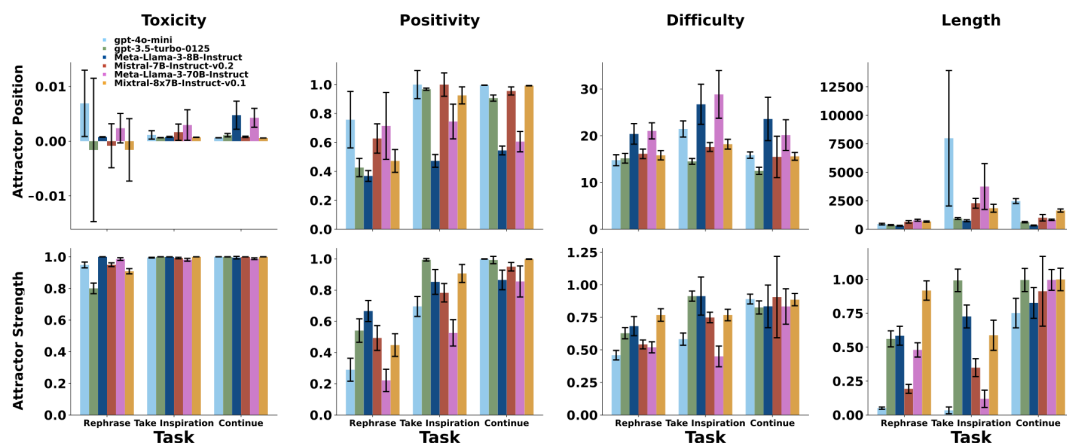


Figure 18: Attractors strength and position.

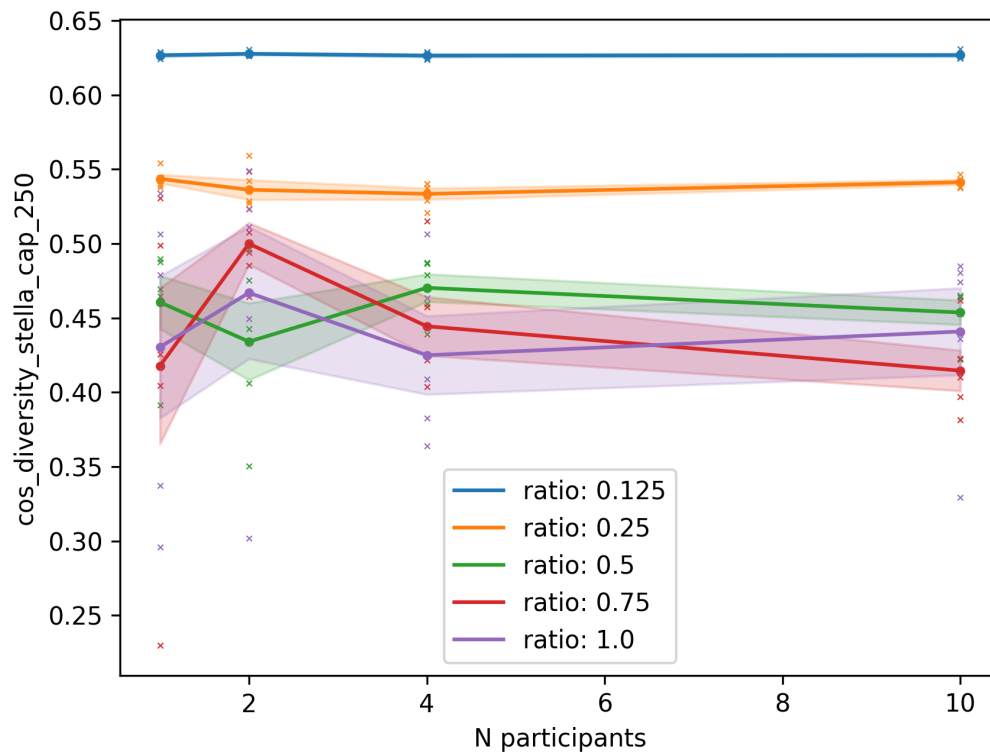
features of the data they are trained on. To empirically evaluate whether model collapse is influenced by dataset properties, we ran iterative fine-tuning experiments manipulating various features of the initial data. Specifically, we varied the ratio of synthetic-to-human data, as well as the quality and the diversity of the initial human data. To measure model collapse along several dimensions, we track how repeated fine-tuning influences the quality and diversity of the data generated by the LLMs. We find that while low ratios of synthetic-to-human data tend to prevent model collapse, the effect of this parameter is not linear. We also show that quality has a very noticeable effect on model collapse, leading to very different dynamics on different internet domains (e.g. Political Twitter vs Reddit). Diversity, on the other hand, does not appear to have such a strong effect. Finally, we find that the tendencies observed when running iterative chains are preserved when scaling the number of agents from 20 to 400, which suggests that using this simplified setting is indeed suited for making prediction about real-world dynamics. Overall, our work extends the existing literature on model collapse by showing that this phenomenon is highly dependent on the human data on which training occurs. This predicts that we might observe collapse at different paces and along different dimensions for different internet domains.

### 8.2.7 Inferring the Phylogeny of Large Language Models and Predicting their Performances in Benchmarks

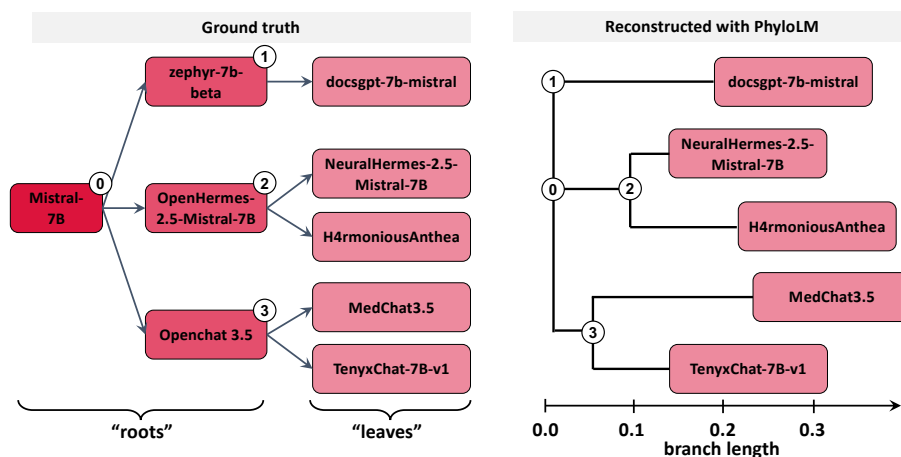
**Participants:** Nicolas Yax (*correspondant*), Pierre-Yves Oudeyer, Stefano Palminteri.

In recent month the number of Large Language Models (LLMs) released has never been that high. On one hand, multiple private companies such as OPENAI, Claude, Google, Mistral, etc. are making cutting-edge models that have a lot of visibility in our modern society and science. However, as the number of LLMs is raising, the training methods are becoming more secretive making the field increasingly obscure to science. On the other hand, everyday, a few hundreds of open-access language models are uploaded on the hugging face hub which is far too much to keep track of the evolution of LLMs in the field. Knowing that not all of these open models are perfectly transparent about the training methods and only very few of them are benchmarked (due to the high cost of benchmarking) there is an increasing need for methods to help keep track of the progress and evolution of these models in the field.

We developed an algorithm inspired from phylogenetics to compute evolutionary trees in LLMs. We show this method efficient in reconstructing the evolutionary history of LLMs within families [20](#), in discriminating the different families, and also in finding similarities between these families. Additionally, the genetic information can be used to predict LLM capabilities like benchmark scores showing a very significant correlation between predicted and true scores. These advances could be instrumental in our



**Figure 19:** The diversity of texts generated at the last (20th) generation as a function of the number of agents per generation. We see that chains with higher percentages of AI data converge to lower diversity text (collapse more). Furthermore, that conclusion generalizes to chains with much more agents.



**Figure 20:** Phylogenetic tree reconstruction. On the left it is shown the ground truth concerning the relation of some LLMs of the Mistral family. Right is the reconstruction from the phylogenetic algorithm for the five latest models of this family ("leaves" of the phylogenetic tree) on which we run PhyloLM. On the right, it is shown the reconstructed phylogenetic tree PhyloLM on the 5 "leaves" models. The numerical labels (0:3) map the true common ancestors (on the right, "ground truth") to the inferred ones (on the left, "reconstructed"). It can be seen that the true and the reconstructed trees are topologically equivalent

way to navigate the field of LLMs by making the world of LLM more transparent at a very low cost.

### 8.3 An Eco-Evo-Devo perspective on Artificial Intelligence

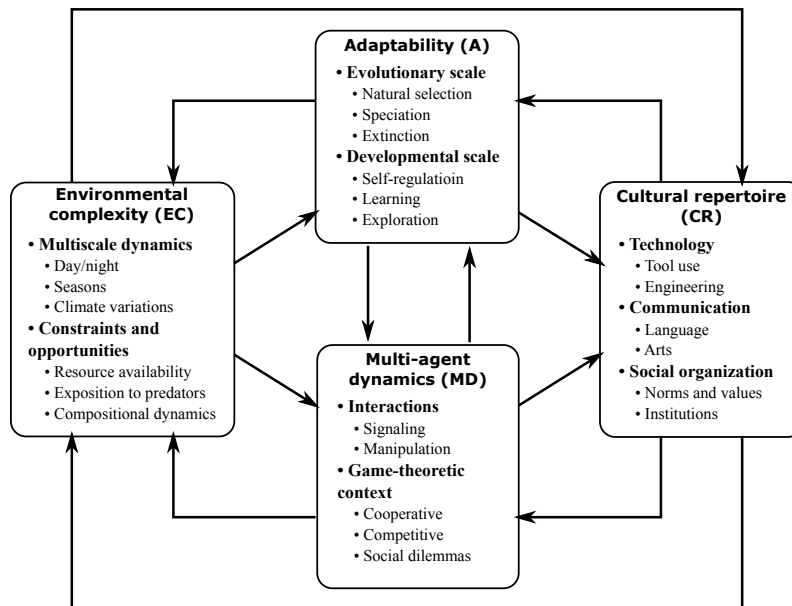
#### 8.3.1 Research perspective: The Ecology of Open-Ended skill Acquisition

**Participants:** Clément Moulin-Frier (*correspondant*), Eleni Nisioti, Pierre-Yves Oudeyer.

An intriguing feature of the human species is our ability to continuously invent new problems and to proactively acquiring new skills in order to solve them: what is called Open-Ended Skill Acquisition (OESA). Understanding the mechanisms underlying OESA is an important scientific challenge in both cognitive science (e.g. by studying infant cognitive development) and in artificial intelligence (aiming at computational architectures capable of open-ended learning). Both fields, however, mostly focus on cognitive and social mechanisms at the scale of an individual's life. It is rarely acknowledged that OESA, an ability that is fundamentally related to the characteristics of human intelligence, has been necessarily shaped by ecological, evolutionary and cultural mechanisms interacting at multiple spatiotemporal scales.

We have recently initiated a new research direction aiming at understanding, modeling and simulating the dynamics of OESA in artificial systems, grounded in theories studying its eco-evolutionary bases in the human species. For this aim, we have proposed a conceptual framework, called ORIGINS (illustrated Fig. 21 and developed in [149]), expressing the complex interactions between environmental, adaptive, multi-agent and cultural dynamics. This framework raises three main research questions:

- What are the ecological conditions favoring the evolution of autotelic agents?
- How to bootstrap the formation of a cultural repertoire in populations of adaptive agents?



**Figure 21:** The ORIGINS framework identifies central components (boxes) and their interactions (arrows) driving Open-Ended Skill Acquisition, both in terms of its evolution from environmental complexity (roughly: left to right arrows) as well its open-ended aspect through feedback mechanisms (right to left arrows). The employed terminology reflects a diversity of mechanisms considered in both Artificial Intelligence and Human Behavioral Ecology.

- What is the role of cultural feedback effects in the open-ended dynamics of human skill acquisition?

The contributions described below are addressing some aspects of these research questions. Note that there might be a thematic overlap between the two last research questions outlined above and the previous section on Models of Cultural Evolution 8.2, where we also present related results.

### 8.3.2 Eco-evolutionary Dynamics of Non-episodic Neuroevolution in Large Multi-agent Environments

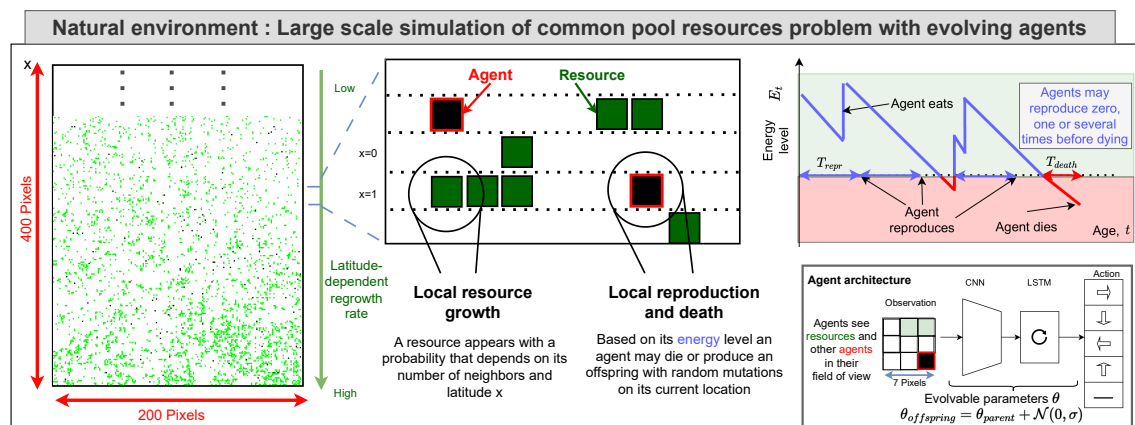
**Participants:** Gautier Hamon (*correspondant*), Eleni Nisioti, Clément Moulin-Frier.

This work was published in 2023 but we keep it in this report as it introduces a general computational framework, called non-episodic neuroevolution, that forms the basis of the two next contributions.

This contribution focuses on eco-evolutionary dynamics where "organisms are not solely products but, by modifying their niche and therefore its associated fitness landscape, are also causes of evolution" [134]. The main objective of this paper is to propose a method for studying large-scale eco-evolutionary dynamics in agent-based simulations with a reasonable level of biological and ecological plausibility. For this aim, we implement a system with the following properties (see Fig. 22 for illustration):

- *Non-episodic simulation environment with complex intrinsic dynamics.* We model our environment after common-pool resource (CPR) appropriation problems, where a group of agents competes for finite resources. We extend an existing environment of CPR appropriation [163] with the presence of multiple niches, where resources regrow proportionally to the density of nearby resources at different rates in different regions of the environment (Fig 22). We prevent any environment or population reset during a whole simulation run, enabling coupled environmental and population dynamics leading to complex eco-evolutionary feedback effects.
- *Continuous neuroevolution in a large, size-varying agent population* The environment contains thousands of agents, each controlled by a neural network whose weights are optimized using neuroevolution [179]





**Figure 22:** Our simulation environment (Left) is an extension of the Common Pool Resource (CPR) environment [163, 136] : a two-dimensional grid-world where some cells contain resources (in green) that the agents (in black) can collect. Resources grow depending on the presence of other resources around them (local growth, Middle) with an additional very sparse spontaneous growth, which means that over-consumption may lead to their local depletion. We introduce a latitudinal model of resource regrowth. We prevent any environment and population reset during a whole simulation, enabling continual eco-evolutionary dynamics to take place. Each agent may reproduce or die according to a physiological model modulating its energy level as a function of life time and resource consumption (Top-Right). The population size varies during the simulation according to the current amount of available resources and the current ability of agents to collect them. Evolution occurs through the mutation of a parent's network weights when it produces an offspring.

- *Physiology-driven death and reproduction* There is no notion of rewards, agents are instead equipped with a physiological system modulating their energy level according to the resources they consume, in a non-linear way. At the evolutionary scale, agents reproduce as long as they are able to maintain their energy level within a reasonable range and die if this level goes below a minimum threshold. This is departure from the notion of fitness-based selection and more in line with a minimal criterion selection [92]. Note that the population size can vary with time.

In addition to experiments conducted in the large environment presented, we also conduct experiments in "lab environment" (as opposed to the "natural environment") to isolate the study of certain behavior (which are often intertwined with a lot of dynamics in the natural environment).

One interesting results of these simulation is the emergence of sustainable foragers which as shown in lab environment Fig.23 tends to not overconsume when there is enough resource in their neighbourhood. This allows to keep a certain amount of resource to spread which is therefore beneficial for their future survival as well as the survival of their offspring. (as there is no reset of the environment)

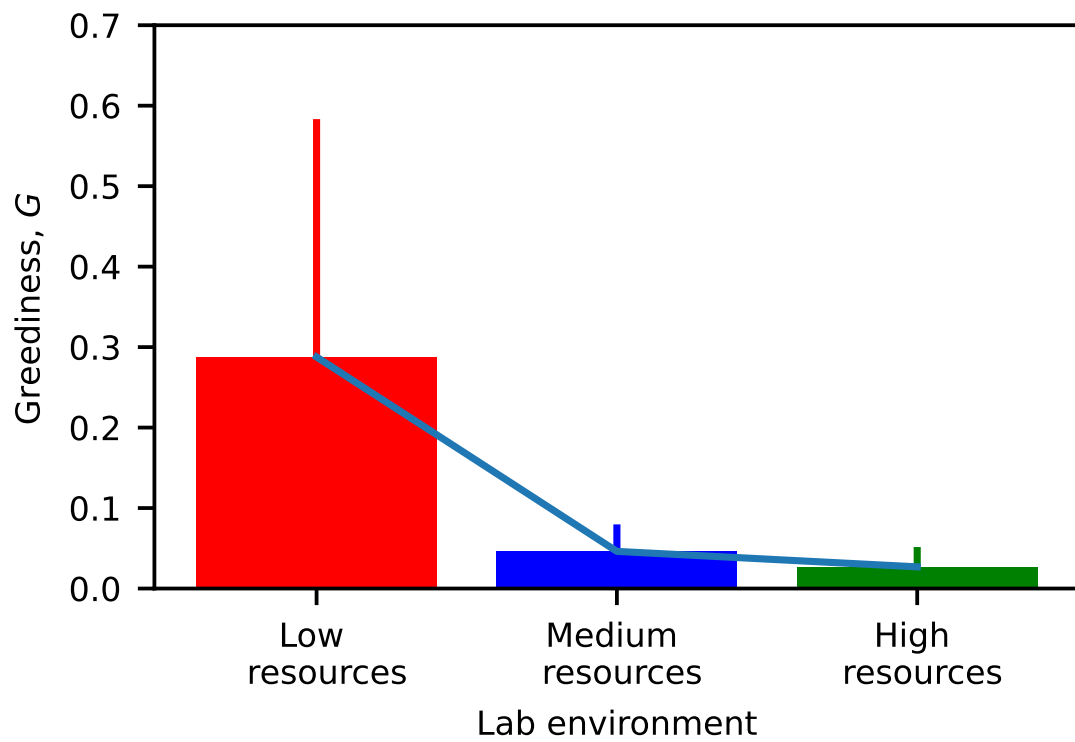
This work was published at the Genetic and Evolutionary Computation Conference (GECCO) 2023. The computational framework it introduced led to the two next recent contributions.

### 8.3.3 Emergent kin selection of altruistic feeding via non-episodic neuroevolution

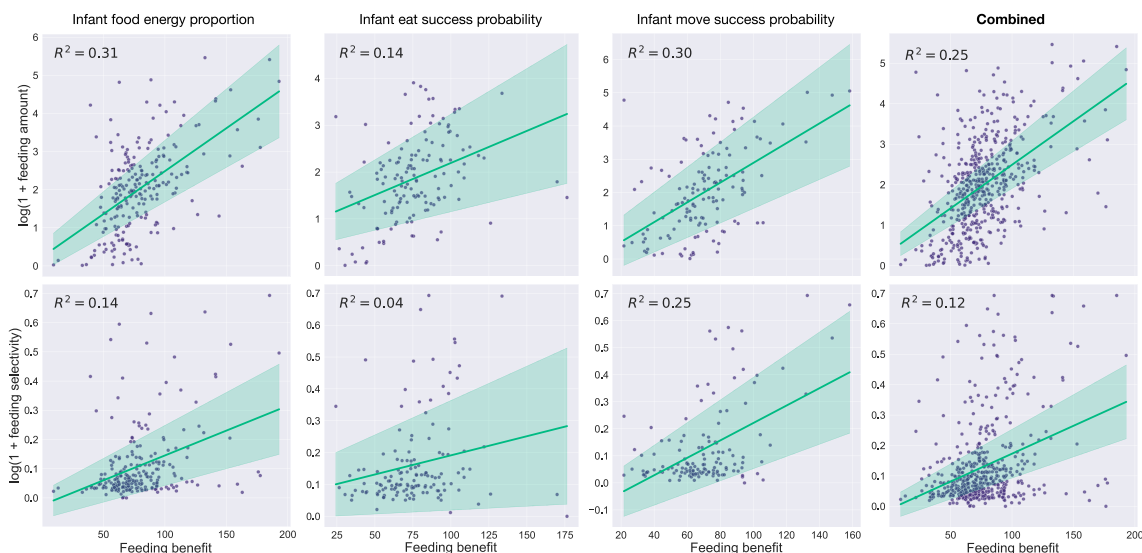
**Participants:** Max Taylor-Davies, Gautier Hamon, Timothe Boulet, Clément Moulin-Frier (*correspondant*).

This work extends the project presented in previous contribution Sec.8.3.2. It is the result from the visit in the team of Max Taylor-Davies doing his PhD at School of Informatics, University of Edinburgh, Scotland. It has been accepted at the EvoStar conference the International Conference on the Applications of Evolutionary Computation (Part of EvoStar) [54].

At first glance, it seems difficult to square the phenomenon of purely altruistic behaviour (acts which confer a benefit to the recipient at a cost to the actor) with the basic principle of natural selection: how can a gene be selected for when it decreases, rather than increases, the fitness of its host? One plausible



**Figure 23:** Greediness of a sustainable forager agent across evaluation environments that differ in the amount of resources. Sustainable agents are far less greedy in environments where there is a certain amount of resources available. This strategy allows to keep resources so that they spread and avoid overdepletion of resources.



**Figure 24:** The relationship between the estimated benefit to infants of being fed and both the amount and selectivity of feeding observed, shown separately for each of the three experimental parameters we varied (and combined in the rightmost column). Each scatterplot point represents a single 500k-timestep simulation run (with values averaged over the final 50k timesteps); regression lines (with 95% confidence intervals) are shown in green. Note that the y-axis shows log(measure) for both amount and selectivity.

account can be made through the theory of inclusive fitness. Key to this theory is the recognition that individual organisms within a social environment are not isolated from their conspecifics in terms of fitness. Whether a given gene is selected for is thus determined by its effect(s) on the fitness of any bearers of copies of that gene. Under this view, we can think of an altruistic act as an exchange of fitness from one agent to another. If the exchange is positive-sum and both sides are bearers of the gene in question, then from the gene's perspective the behaviour confers a fitness benefit—even while it decreases the fitness of the acting individual.

Kin selection theory [178] has proven to be a popular and widely accepted account of how altruistic behaviour can evolve under natural selection. Hamilton's rule, first published in 1964 [118, 119], has since been experimentally validated across a range of different species and social behaviours. In contrast to this large body of work in natural populations, however, there has been relatively little study of kin selection *in silico*. In the current work, we offer what is to our knowledge the first demonstration of kin selection emerging naturally within a population of agents undergoing continuous neuroevolution. Specifically, we find that zero-sum transfer of resources from parents to their infant offspring evolves through kin selection in environments where it is hard for offspring to survive alone. In an additional experiment, we show that kin selection in our simulations relies on a combination of kin recognition and population viscosity. We believe that our work may contribute to the understanding of kin selection in minimal evolutionary systems, without explicit notions of genes and fitness maximisation.

This paper was accepted at The International Conference on the Applications of Evolutionary Computation (EvoAPPS) 2025 (part of EvoStar).

### 8.3.4 Evolving large populations of adaptive neural agents in ecologically plausible environments

**Participants:** Timothé Boulet (*correspondant*), Gautier Hamon, Clément Moulin-Frier.

This work continues the project presented in the previous paragraph, with a focus on the ability of agents to develop adaptability behaviors. Specifically, we extend the framework by adding fruits, a spatially variable resource, and a memory of the values of each type of fruits for the agents. The goal is to

observe whether the agents manage to exploit the knowledge of the fruits values to decide which fruit exploit.

Results : the agent were able to exploit the fruit value information to optimize their behavior. There were also some results that we were not necessarily expecting and that comes from our choice of model. Notably, it seems the agents choice for exploiting a cluster is heavily influenced by social criteria (the number of agents already exploiting it) and cultural criteria (whether the cluster is empty or full of fruits). This effect exceeds the adaptability effect in the latest stages of the simulations.

### 8.3.5 Cooperative control of environmental extremes by artificial intelligent agents

**Participants:** Martí Sànchez Fibla, Clément Moulin-Frier (*correspondant*), Ricard Solé.

This contribution is the result of a collaboration between Ricard Solé (University Pompeu Fabra (Barcelona, Spain) and Martí Sànchez-Fibla (CSIC, Barcelona, Spain) and Clément Moulin-Frier (Flowers, Inria). It has been published in the Journal of the Royal Society Interface in 2024 [39].

Humans have been able to tackle biosphere complexities by acting as ecosystem engineers, profoundly changing the flows of matter, energy and information. This includes major innovations that allowed to reduce and control the impact of extreme events. Modelling the evolution of such adaptive dynamics can be challenging given the potentially large number of individual and environmental variables involved. This paper shows how to address this problem by using fire as the source of external, bursting and wide fluctuations. Fire propagates on a spatial landscape where a group of agents harvest and exploit trees while avoiding the damaging effects of fire spreading. The agents need to solve a conflict to reach a group-level optimal state: while tree harvesting reduces the propagation of fires, it also reduces the availability of resources provided by trees. It is shown that the system displays two major evolutionary innovations that end up in an ecological engineering strategy that favours high biomass along with the suppression of large fires. The implications for potential A.I. management of complex ecosystems are discussed.

The computational model is illustrated Fig. 25 and the main results in Fig. 26.

### 8.3.6 Emergence of agriculture in an artificial society of reinforcement learning agents

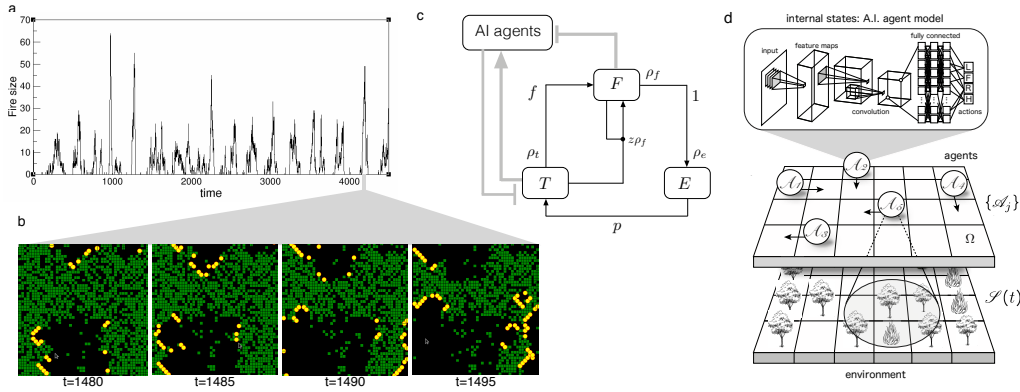
**Participants:** Gautier Hamon (*correspondant*), Clément Moulin-Frier, Marti Sanchez-Fibla, Ricard Solé.

This work is still a work in progress. It is the result of a collaboration with Ricard Solé (Complex System Lab, Universitat Pompeu Fabra, Barcelona, Spain) and Martí Sànchez-Fibla (CSIC, Barcelona, Spain), with a 3 months visit of Gautier Hamon in the Complex system lab.

This project aims at exploring the emergence of agriculture in simulations of artificial agents. In particular, we study how group of agents trained independently with reinforcement learning can learn to collectively engineer their environment to improve the growth of a specific beneficial plant that they can eat. We explore the environmental and agentic conditions that favor the emergence of agriculture in the simulations.

Our simulated environment consists in a 2D map with discrete locations (a lattice). We simulate the growth of different types of competing plants on this map, with only one being beneficial to the agent. We simulate plants lifecycle with the growth of the plant and the local spread of seeds. The environment also contains sources of water. The presence of water at the same location of a plant speed up its growing, i.e. water acts as a fertilizer.

This dynamic allows the agent to potentially eco-engineer their environment through the removal of unwanted plants, spreading seeds or watering the plants. We observe that in some environmental and agentic conditions, agents effectively learn to collectively improve the growth of the beneficial plant, allowing for an increase in resources to eat. Interestingly the agents learn a collective strategy to not



**Figure 25:** Forest fire dynamics in time and space. In figure (a) a typical time series of the number of burned sites in a forest fire model (FFM) is displayed for a square lattice with  $L = 50$  and parameters  $p = 0.003$ ,  $f = 0.00003$ . The number of sites burning (the fire size) shows marked bursting dynamics. Four spatial snapshots are shown in (b) associated with a fire burst. Here, green, yellow and black correspond to trees, fires and ashes (empty sites) respectively. The basic set of rules is summarized in (c) using black arrows. In our model, we add a set of AI agents whose interactions with the environment are marked with grey arrows where positive and negative interactions are indicated as  $\rightarrow$  and  $\leftarrow$ , respectively. They benefit from trees but get punished by fire spreading, and can modify tree density by harvesting trees. In (d) we summarize the levels of interaction between forest fire dynamics and its control by neural agents. The bottom layer defines the observed spatial pattern of states of the Forest Fire Model (FFM), which changes stochastically while can be affected by the action of agents (middle layer) that have a limited observation range (indicated as a circle in the bottom layer) and can take decisions about their movement and harvesting trees locally. Each agent (upper layer) makes decisions (implements an action policy, mapping observed states to actions) by means of a convolutional neural network trained with Reinforcement Learning (RL). The RL process eventually defines the behavioural pattern displayed by the agent, which translates into a set of potential actions (*LeftTurn*, *RightTurn*, *ForwardMove*, *Harvest*) in response to the local environment.

overconsume to allow further spreading of the beneficial plant. In addition, we observe specialization, with some agents planting more seeds while other bring water or remove the unwanted plants.

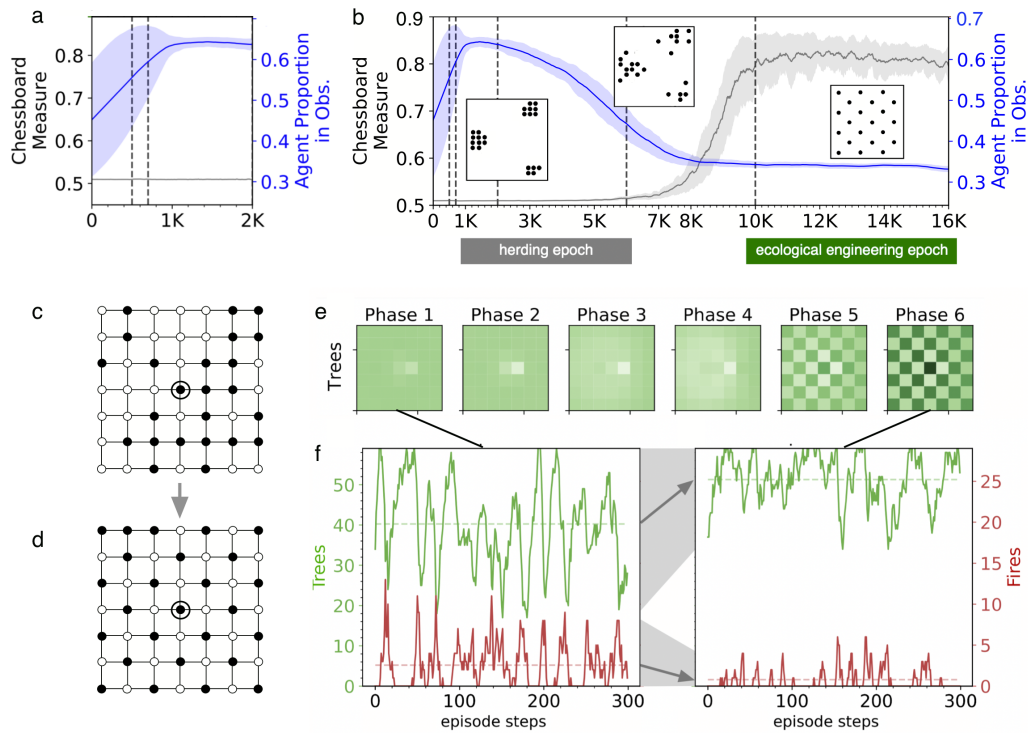
Also, in a setup where agent can reproduce if they have enough resources, we show that the emergence of "agriculture" (eco-engineering the environment to have more abundant sustained resources), comes with a seemingly increase in population size. These new individuals also adopt an "agriculture" behavior and allow to grow even more beneficial plant through their workforce, allowing again to grow the population.

These simulations therefore could help understand some of the dynamics underlying the emergence of agriculture. Further work need to be done to confirm and analyse more thoroughly the results that we obtained in these simulations. This work can also benefit the multi-agent RL community with the emergence of strong collective capabilities.

### 8.3.7 Open-ended recipe crafting through meta reinforcement learning

**Participants:** Gautier Hamon (*correspondant*), Eleni Nisioti, Clément Moulin-Frier.

As a first step towards studying the evolution of open-ended skill acquisition in artificial agents, we studied the environmental conditions favoring the systematic exploration of combinatorial recipes involving various objects. By *combinatorial recipe*, we mean the ability of agents to combine objects in the environment in order to create new ones (in the spirit of the Minecraft video game), some of these crafted objects being associated with a reward. In this work, the training of an agent uses *meta reinforcement learning* where an outer loop, equivalent to an evolutionary mechanism, meta-learns the parameters of an inner loop which can be seen as a developmental mechanism (where the agent acquire skills during its lifetime by interacting with the environment).



**Figure 26:** Characterization of the observed cooperative phase transitions. In (a-b), the time series for the herding (blue) and ecological-engineering (grey) measures are displayed. The herding measure characterizes the agents' tendency to form dense herds. It is illustrated in the three insets in (b), sketching typical spatial arrangements of agents: a high herding measure indicates that agents form dense herds (left inset); in contrast, lower herding measures indicate that they are more uniformly spread in the environment (middle and right inset). The ecological-engineering measure characterizes the agents' ability to create a structured pattern of trees limiting fire propagation. Patterns of trees corresponding to a low and high ecological-engineering measure are illustrated in (c) and (d), respectively. The open circle in the middle indicates an agent, with black and white circles indicating the presence or absence of trees, respectively. Intuitively, the high measure resulting from (d) corresponds to a perfect chessboard pattern preventing fire propagation (which only propagates in the horizontal and vertical dimensions) while maximizing the number of trees. The progressive formation of this structured pattern of trees is displayed in (e), showing the average density of trees in all agent's neighbourhoods during the six emerging phases. In (f), we observe the FFM dynamics in the first and last episodes of the simulation. This demonstrates that the agent population managed to increase (resp. decrease) the average number of trees (resp. fire) as indicated by the horizontal dotted lines while reducing the fluctuation range of trees and fires.

In the current setup we use  $RL^2$  as our meta-learning algorithm [112, 188] which has already been used for acquiring behaviors efficiently balancing exploration and exploitation in a simple navigation task. Other work studied how different conditions in a bandit favor the evolution of innate vs learned behaviors [135].

Our experiments with recipe crafting are inspired by [the little alchemy game](#). The difference with previous works in similar environments (e.g. 8.2.1) is that at every episode the structure of the recipe is randomly chosen. The agent therefore cannot predict what recipes will be rewarding and have to explore different combinations of objects in order to find the rewarding ones. The agent should also memorize the successful and unsuccessful combinations in order to explore and exploit efficiently.

Our preliminary results use both in a vectorized version of the game (where the agents action are only to choose the 2 objects to combine) and an embodied gridworld version (where the agent has to move, grab objects and put them on top of others in order to craft new ones). In both of these cases, the training efficiently meta learns an exploration/exploitation strategy which is to try new recipes (most of the time it does not try non working recipes more than once) until it finds the rewarding ones and then simply exploits them by making them over and over.

In preliminary results, by introducing drastic changes of recipe during an episode we observed the meta-learning of curious exploration just for the sake of information. Agents despite not being rewarded for it, explore the possible recipe as a way to get information that might be useful later in case of the change.

Further work will explore how increasing the variability (potentially through a curriculum) could lead to never ending exploration of the environment. In addition, adding other types of environmental dynamic to explore might lead to general curious behavior.

During the project, we switched from evolutionary algorithm and Recurrent neural networks to reinforcement learning and transformers. This allowed for more complex environments with more possibilities. We also obtained preliminary results with agents exploring the environment to gain information for the future.

This work uses the [JAX python library](#) for both the model/machine learning part and the environment simulation. JAX allows easy parallelization and fast GPU computation and so learning it through this project will be useful for later projects.

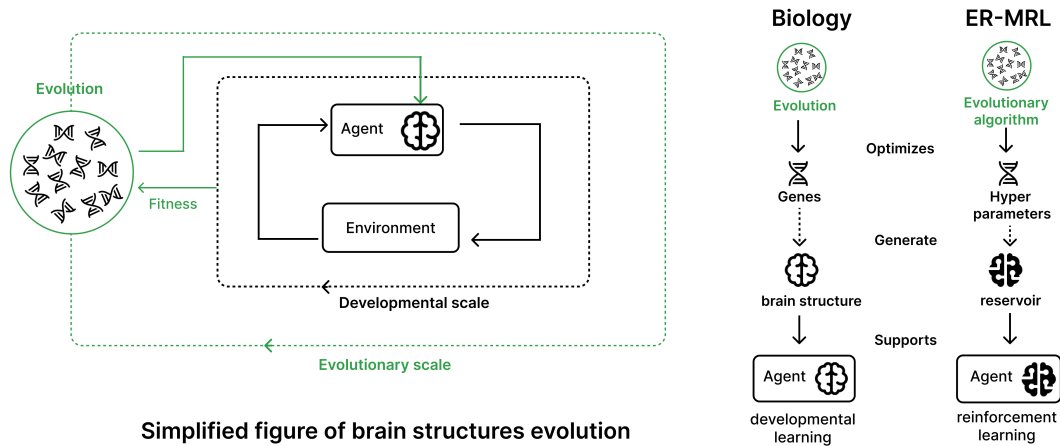
### 8.3.8 Evolving Reservoirs for Meta Reinforcement Learning

**Participants:** Corentin Leger (*correspondant*), Gautier Hamon, Eleni Nisioti, Xavier Hinaut, Clément Moulin-Frier.

This contribution was realized in the context of the Master internship of Corentin Léger in 2023, as a collaboration between C. Moulin-Frier from the Flowers team and Xavier Hinaut from the Mnemosyne team. It led to a paper which has been published in the International Conference on the Applications of Evolutionary Computation (Part of EvoStar) [48] in April 2024.

Animals demonstrate remarkable adaptability to their environments, a trait honed through the evolution of their morphological and neural structures [184][162]. Animals are born equipped with both hard-wired behavior routines (e.g. breathing, motor babbling) and learning capabilities to adapt from experience. The costs and benefits of evolving hard-wired behaviors vs. learning capabilities depends on different factors, a central one being the level of unpredictability of environmental conditions across generations [181][130]. Phenotypic traits addressing environmental challenges that are shared across many generations are more likely to evolve hard-wired (e.g. breathing), while traits whose utility can hardly be predicted from its utility in previous generations are likely to be learned through individual development (e.g. learning a specific language).

This prompts an intriguing question: How can neural structures, optimized at an evolutionary scale, enhance the capabilities of agents to learn complex tasks at a developmental scale? To address this question, we propose to model the interplay between evolution and development as two nested adaptive loops: evolution optimizes the generation of neural structures through natural selection over generations, shaping developmental learning during an agent's lifetime (Fig. 27).



**Figure 27:** A simplified view of the evolution of brain structures (left) and the parallel with our computational approach (right). We can observe on the left of the figure the interplay between two loops : an evolutionary one that modifies the generating parameters of neural structures, and a developmental one where agents equipped with such neural structures learn to interact with their environment. We propose a computational framework (right) where an evolutionary algorithm optimizes hyperparameters that generate neural structures called reservoirs. These reservoirs are then integrated into RL agents that learn an action policy to maximize their reward in an environment

More precisely, at the evolutionary scale (the outer loop), we use an evolutionary algorithm to optimize a genome specifying Hyper Parameters of Reservoirs [173]. At a developmental scale (the inner loop), a RL agent equipped with a generated reservoir learns an action policy to maximize cumulative reward in a simulated environment. Thus, the objective of the outer evolutionary loop is to optimize macro properties of reservoirs in order to facilitate the learning of an action policy in the inner developmental loop. See Fig.28 for an overview of the method.

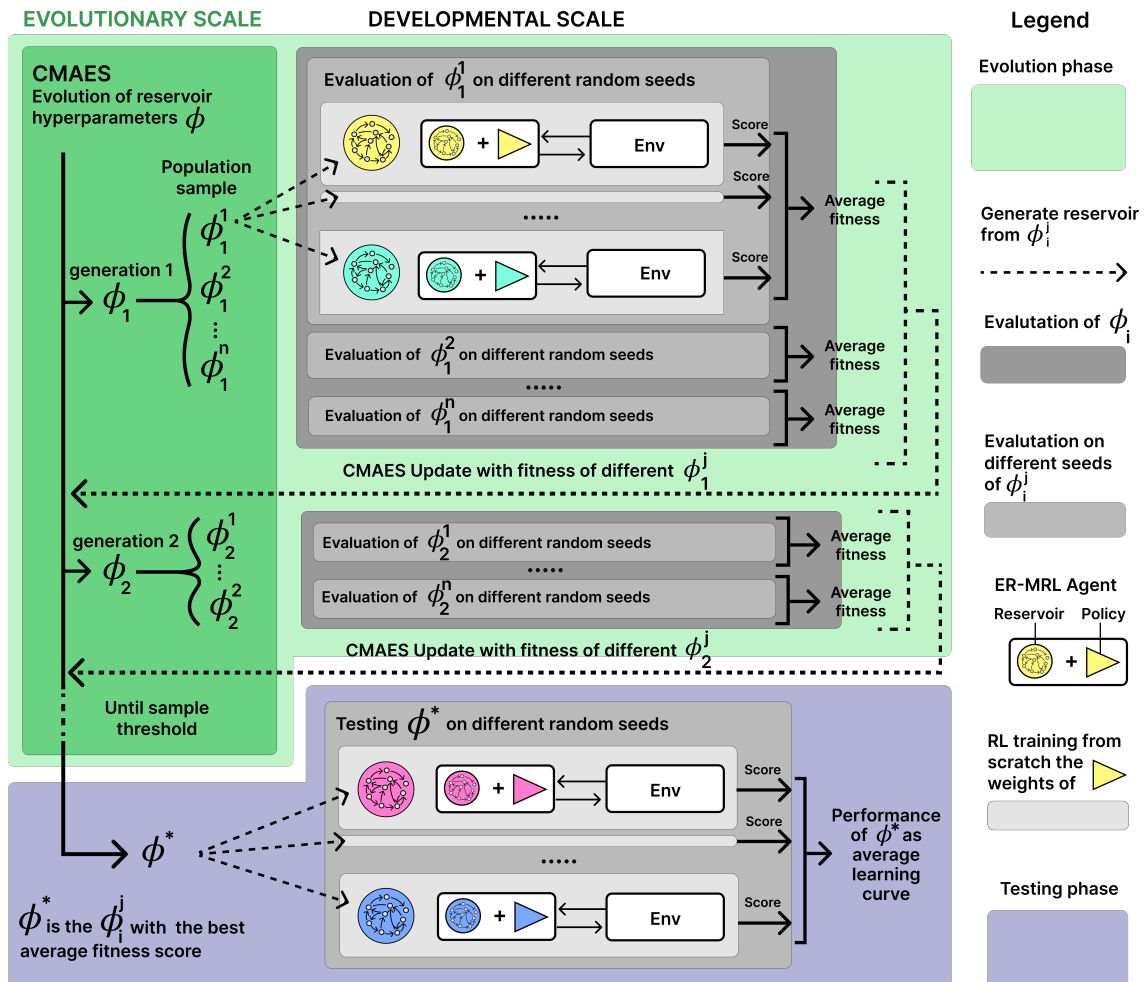
Using this computational model, we run experiments in diverse simulated environments, e.g. 2D environments where the agent learns how to balance a pendulum and 3D environments where the agent learns how to control complex morphologies. These experiments provide support to three main hypotheses for how evolved reservoirs can affect intralife learning. First, they can facilitate solving partially-observable tasks, where the agent lacks access to all the information necessary to solve the task. In this case, we test the hypothesis that the recurrent nature of the reservoir will enable learning to infer the unobservable information. Second, it can generate oscillatory dynamics useful for solving locomotion tasks. In this case, the reservoir acts as a meta-learned CPG. Third, it can facilitate the generalization of learned behaviors to new tasks unknown during the evolution phase.

## 8.4 Theories of human curiosity-driven learning

**Participants:** Gaia Molinari (*correspondant*), Pierre-Yves Oudeyer, Cedric Colas, Anne Collins.

Humans are autotelic agents who learn by setting and pursuing their own goals. However, the precise mechanisms guiding human goal selection remain unclear. Learning progress, typically measured as the observed change in performance, can provide a valuable signal for goal selection in both humans and artificial agents. In a new paper published at Neurips 2024 [50], we hypothesize that human choices of goals may also be driven by latent learning progress, which humans can estimate through knowledge of their actions and the environment – even without experiencing immediate changes in performance. To test this hypothesis, we designed a hierarchical reinforcement learning task in which human participants (N = 175) repeatedly chose their own goals and learned goal-conditioned policies. Our behavioral and computational modeling results confirm the influence of latent learning progress on goal selection and





**Figure 28:** A simplified view of the evolution of brain structures (left) and the parallel with our computational approach (right). We can observe on the left of the figure the interplay between two loops : an evolutionary one that modifies the generating parameters of neural structures, and a developmental one where agents equipped with such neural structures learn to interact with their environment. We propose a computational framework (right) where an evolutionary algorithm optimizes hyperparameters that generate neural structures called reservoirs. These reservoirs are then integrated into RL agents that learn an action policy to maximize their reward in an environment

uncover inter-individual differences, partially mediated by recognition of the task's hierarchical structure. By investigating the role of latent learning progress in human goal selection, we pave the way for more effective and personalized learning experiences as well as the advancement of more human-like autotelic machines.

## 8.5 Generative AI and educational technologies

### 8.5.1 Fostering curiosity and meta-cognition in children using LLM-based conversational agents

**Participants:** Pierre-Yves Oudeyer, H el ene Sauz eon (*correspondant*), Rania Abdelghani, Didier Roy, Edith Law, Chlo e Desvaux.

For empowering our conversational agents embedded into the curiosity-driven intervention we built, we leverage from LLMs, as verbal incentives either to curious behaviors (question asking and information search) or to metacognitive states triggering curious behaviors.

**Matching conversational agents and LLM as lever for scaling up field applications.** At the end of 2020, we started an industrial collaboration project with EvidenceB on this topic (CIFRE, contract of Rania Abdelghani validated by the ANRT). The overall objective of the thesis is to propose new educational technologies driven by epistemic curiosity, and allowing children to express themselves more and learn better. To this end, a central question of the work will be to specify the impact of self-questioning aroused by states of curiosity about student performance. Another objective will be to create and study the pedagogical impact of new educational technologies in real situations (schools) promoting an active education of students based on their curiosity. To this end, a web platform called 'Kids Ask' has been designed, developed and tested in three primary schools. The tool offers an interaction with a conversational agent that trains children's abilities to generate curiosity-driven questions and use these questions to explore a learning environment and acquire new knowledge. The results suggest that the configuration helped enhance children's questioning and exploratory behaviors; they also show that learning progress differences in children can be explained by the differences in their curiosity-driven behaviors [83].

Despite showing pedagogical efficiency, the method used in the first study of this PhD is still very limited since it relies on generating curiosity-prompting cues by hand for each educational resource in order to feed the "discussion" with the agent, which can be a very long and costly process. For this reason, a logical follow-up to scale-up and generalize this study was to explore ways to automate the said conversational agents' behaviors in order to facilitate their implementation on a larger scale and for different learning tasks. More particularly, we move towards the natural language processing (NLP) field and the large language models (LLMs) that showed an impressive ability in generating text that resembles the way people write.

In this context, we study using the recent LLM GPT-3 to implement conversational agents that can prompt children's curiosity about a given text-based educational content, by proposing some specific cues. We investigate the validity of this automation method by comparing its impact on children's divergent question-asking skills with respect to the hand-crafted condition we had in our previous work.

In a second step, we explore using GPT-3 to propose a new curiosity-prompting behavior for our agent that aims to better support the children's needs of competence, autonomy and relatedness during the question-asking training. The study was conducted in two primary schools with 75 children aged between 9 and 11, and assigned to one of the three experimental conditions of agent's design: 1. incentive hand-crafted agent ; 2. the incentive GPT-3-driven agent and 3. the open GPT-3 agent (see agent description hereafter).

Our first results suggest the validity of using GPT-3 to facilitate the implementation of curiosity-stimulating learning technologies. Indeed, children's performance was similar between the conditions where they had hand-generated or GPT-3-generated cues. In a second step, we also found that GPT-3 can be efficient in proposing the relevant cues that leave children with more autonomy to express their curiosity [85] (publication in process).

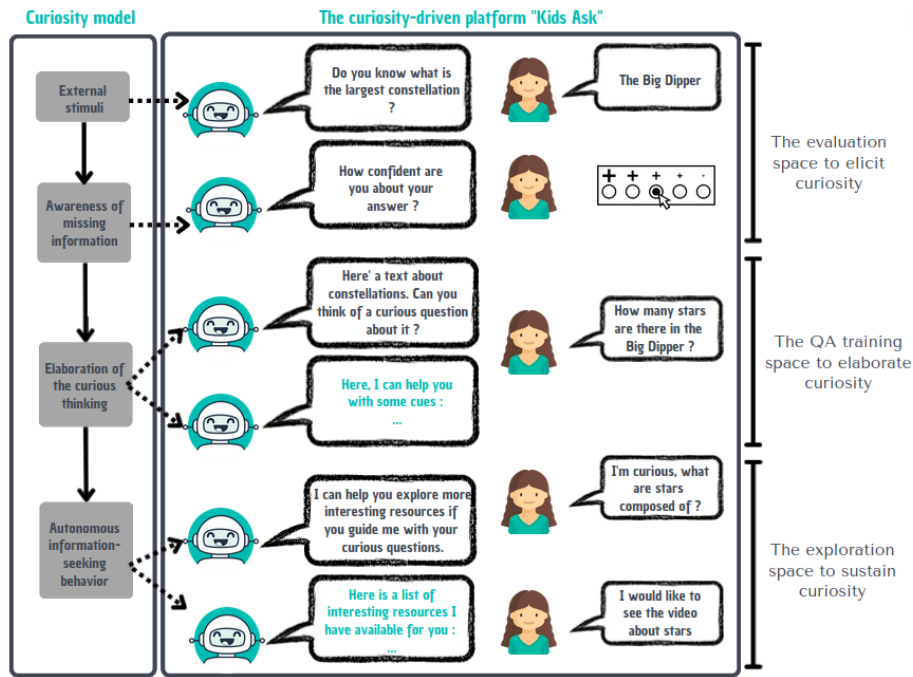


Figure 29: Illustration of a conversational agent’s strategies in the different work spaces of the "Kids Ask" platform

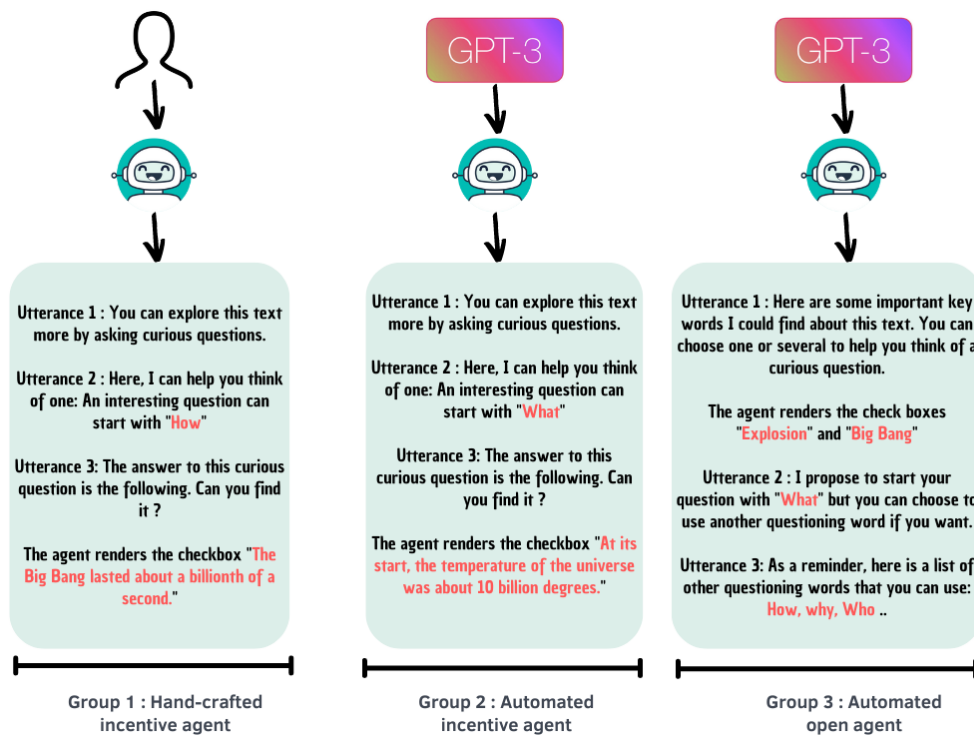
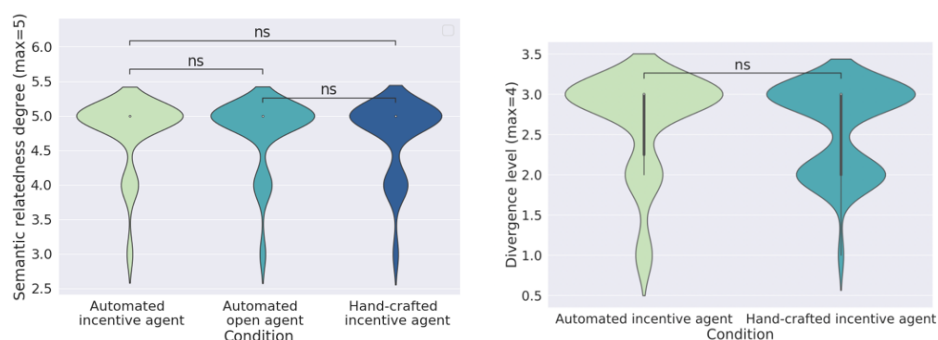


Figure 30: Illustration of utterances for the 3 conditions. From left to right : the incentive hand-crafted agent, the incentive GPT-3-driven agent and the open GPT-3 agent

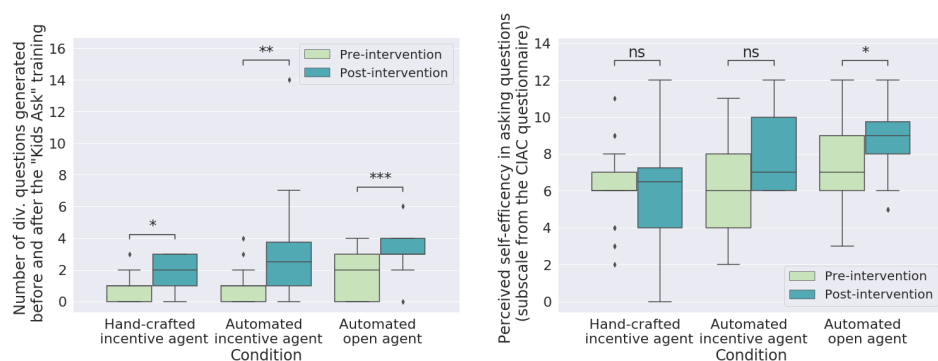
In evaluating the cues generated, we noted no offensive output from GPT3 and the three conditions were similar in terms of both the divergence level and the semantic relatedness to the text with what we

had before by hand (as evaluated by human annotators) (as illustrated in the figure below).



**Figure 31:** Left : Semantic relatedness of the cues from GPT3 is similar to what was produced by hand. Right: Divergence level for the incentive semantic cues was similar for the hand-crafted and the GPT3 versions.

Importantly, the results in terms of learner experience (elicited motivation and cognitive load, as shown below) as well as of curiosity behaviors have been conclusive by revealing equivalent results in children with GPT3-generated incentives, or even a superiority effect of the “open” condition of GPT-3.

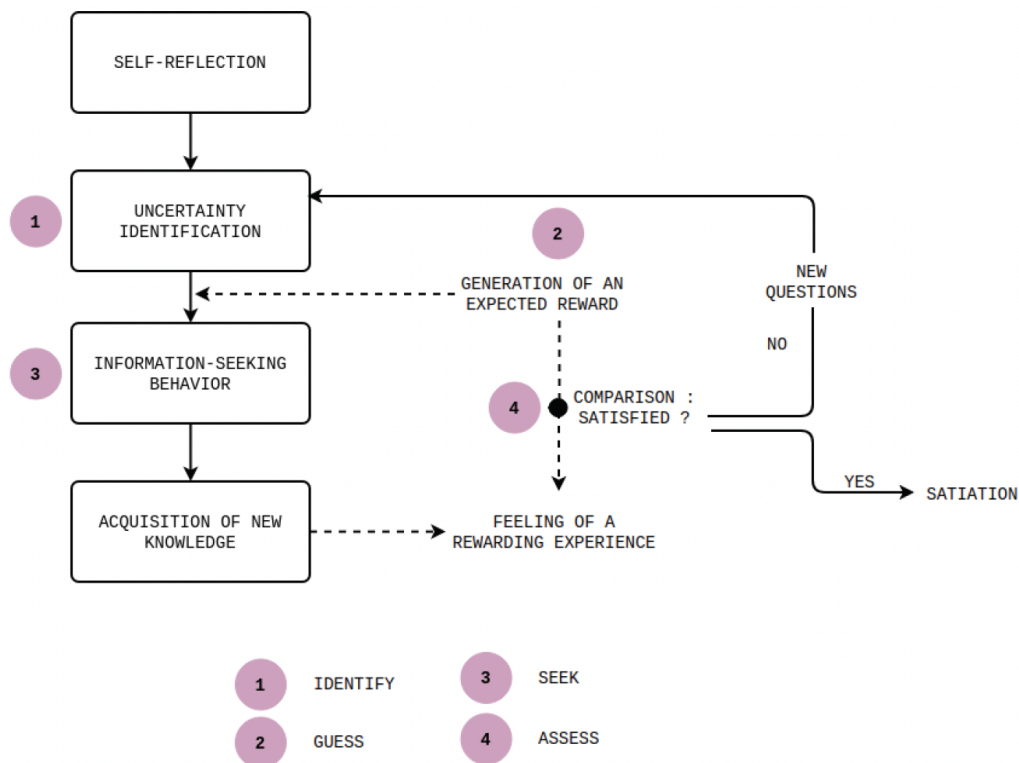


**Figure 32:** **Left:** Participants from the three conditions were able to improve their divergent QA abilities after the “Kids Ask” interaction, as shown by the divergent QA fluency test pre- and post-training. **Right:** Children’s perception of their QA self-efficacy changed more positively with the intervention for those who interacted with the automated agents

### 8.5.2 LLMs usages and curiosity-related metacognitive skills in educational setting

Thanks to a collaboration with Pr. Celeste Kidd (Kidd Lab, Berkeley University of California, US), on another subject, we also started investigating the importance of curiosity-related metacognitive skills on students’ use of the GenAI (Generative AI) tools during learning. Indeed, in [84], we argue about the importance of developing children’s sense of critical thinking, epistemic vigilance, etc in order to allow a more active and informed use of these tools during learning. Such skills can help have a more realistic expectations of such tools and evaluate their outputs before integrating them in one’s beliefs.

**Identify-Guess-Seek-Assess (IGSA), a framework for modeling and training children’s curiosity-driven learning through metacognition.** Leveraging the hypotheses about the roles of metacognition in organizing curiosity-driven learning, we recently introduced a training paradigm aiming to teach curiosity-driven learning via training of metacognitive skills [82]. This training paradigm is based on considering the four basic metacognitive skills of Murayama’s framework [150] : identify (I) a knowledge gap, guess (G) possible answers, seek (S) information to fill the gap, and assess (A) the quality of the information (as shown below).

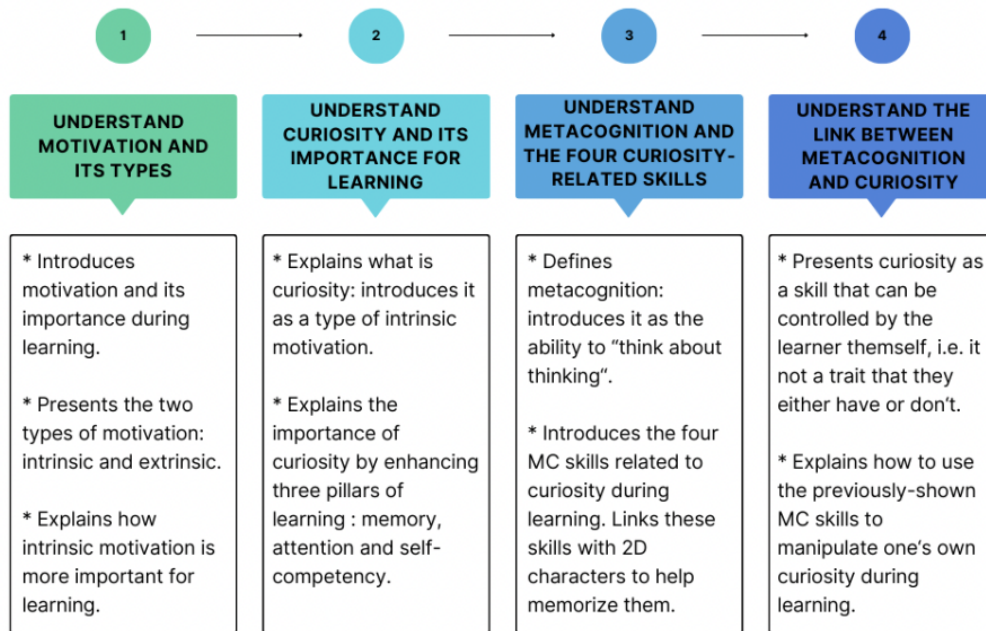


**Figure 33:** Curiosity-driven learning framework and link with the metacognitive skills we propose to train as facilitators during our IGSA-based intervention

**Design.** In the training paradigm, the relevant metacognitive skills are personified as animated verbal metacognitive characters as follows: 1. The first character is the referrer who enacts the “identify” skill: when learning new content, it observes the task, reflects on its previous knowledge and chooses which uncertainty or missing information to pursue. 2. The second character is the detective, reflecting the “guess” skill: it formulates educated guesses and makes predictions about the missing information to pursue. 3. The third character is the explorer reflecting the “seek” skill: it pursues uncertainty by asking the relevant questions or by exploring the relevant resources. 4. The final character reflects the “assess” skill: it evaluates whether the inquiry resulted in learning progress with respect to the initial uncertainty.

The training paradigm is divided into two steps. • The first step aims to help children gain declarative knowledge about the said metacognitive skills. It consists of animated videos explaining key concepts related to curiosity and/or metacognition as well as the four metacognitive skills. To facilitate the explaining of these skills, the videos model them with 2D characters that correspond to their role in the curiosity-driven cycle: a first controller for IDENTIFY skill, a detective for the GUESS skill, an explorer for the SEEK skill and a second for the ASSESS skill (see figure below).

• The second step of training aims to help children gain procedural knowledge about the four Metacognitive skills they saw during the videos. For this we create a web-based platform we name “Kids meta-think”, where they are prompted to use these skills appropriately during a reading-comprehension task, using the help of conversational agents that have the same appearance and roles as the 2D characters presented during the videos. The interaction with the agents is designed to be like the following: once participants finish reading a text, the first referee appears (representing the IDENTIFY skill) and guides them into identifying and formulating a knowledge gap. After validating this step, the detective agent appears to guide them into formulating an educated hypothesis about it (GUESS skill). Then the explorer appears (SEEK skill) to help formulate the appropriate divergent question. And finally, once the divergent answer submitted, we display a set of 3 pieces of information that may or may not contain the answer to the question. The second referee thus appears (ASSESS skill) to lead children into reflecting on these pieces of information and decide whether or not they bring them closer to closing



**Figure 34:** The pedagogical goals and content of the four videos (first step of training)

their knowledge gap. The agents had predefined scripted behaviors and did not give any feedback regarding children's inputs. See following figure for a concrete example of the agents' dialogue for a given text.

**Technical implementation of "Kids meta-think".** The behaviors of the four agents in terms of selection of the appropriate prompts as well as proposing a list of options (if requested) are entirely predefined and hand-scripted. Indeed, the agents are connected to a database containing the text resources and, for each one, a list of corresponding prompts relating to the four metacognitive skills, i.e., the utterances for the conversational agents for each skill. These utterances consist of sentences to remind the definition of the skill, its importance and how to use it. Each text is also associated with a list of 3 propositions for each skill. All of these resources have been previously hand-generated by the research team and validated by two teachers as to their pedagogical relevance.

During the interaction, the agent's automaton composes the dialogue utterances in order to include the appropriate cuing strategies (see figure below). We changed the utterances between the texts to avoid repetition in the different agents' dialogue. This implementation has no natural language processing or generative artificial intelligence methods to manipulate the behaviors for the different agents. All data entered by participants was saved in a local database and was only evaluated post-experimentation, during the data

**Intervention Effectiveness.** As a follow-up direction to this line of work, a first study has been conducted with two classrooms to evaluate the accessibility of this new training and the impact on metacognitive efficiency, curiosity-driven question-asking and learning. The first results being rather positive, a larger sample size was recruited to validate them [82].

To go further in the support of effectiveness of the "learning to learn through curiosity", (PhD project Marie-Sarah Desvaux - started in October 2023 University of Bordeaux), we're currently conducted a field study in a larger scale in collaboration with the Académie de Bordeaux and five elementary school teachers of Bordeaux Métropole, for a total of 100 students aged 9 to 11 years and according to the methodological rules of RCT in field setting; this approach is expected in the educational sciences. The objective is twofold : first the intervention was tested when led by teachers, as this reflects the classical school setting. Teachers underwent short training sessions to familiarize themselves with the format

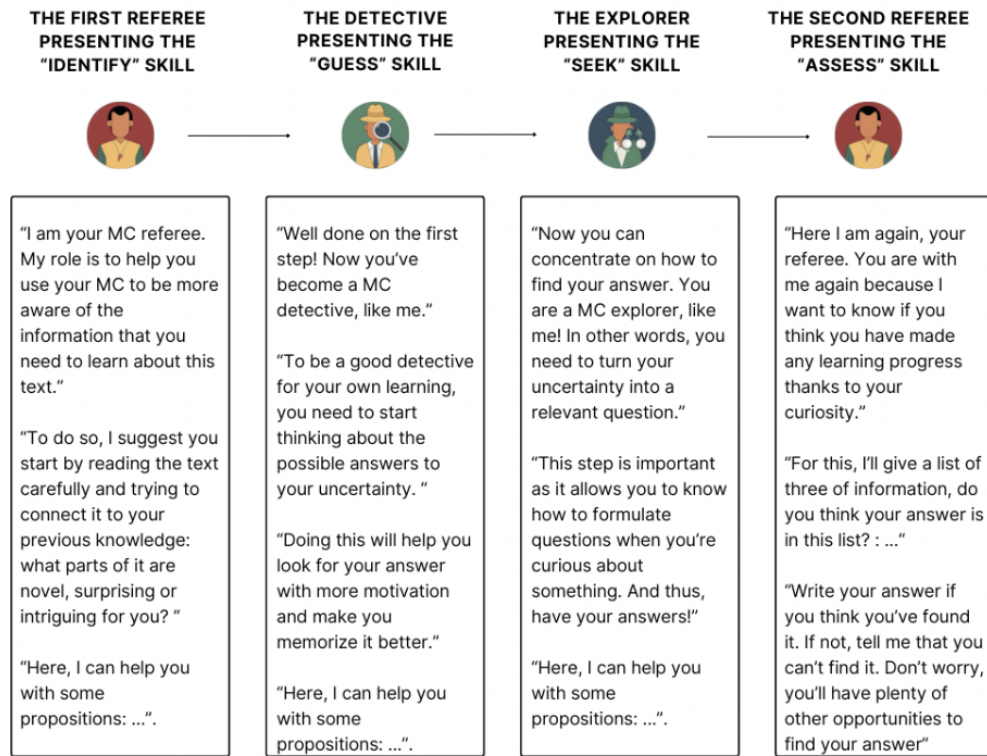


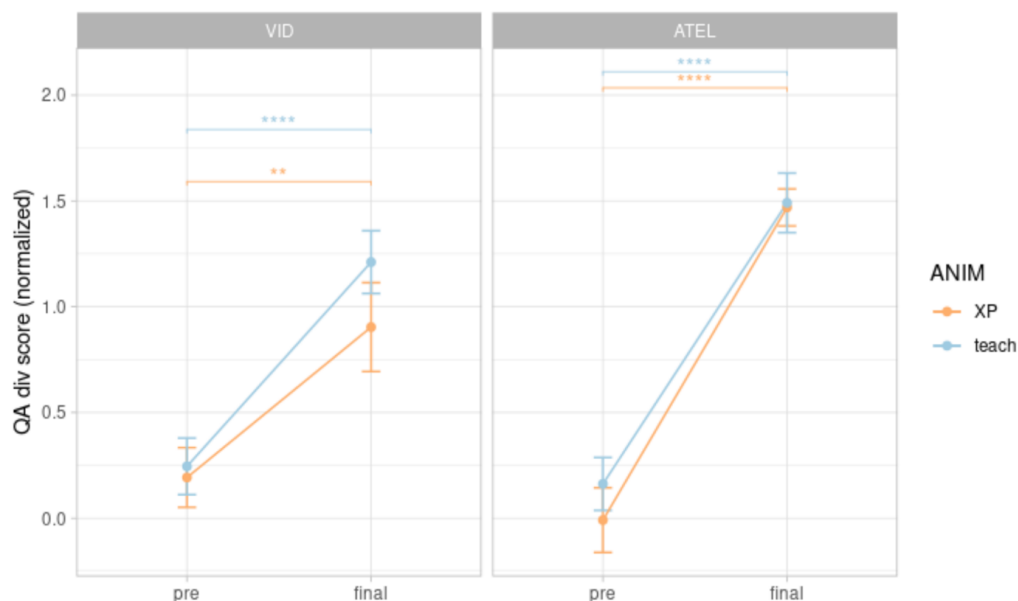
Figure 35: Examples of the four agents' utterances for a given text

Figure 36: Screenshot of the "Kids meta-think" platform during the training, given one text

and content of the intervention, enabling them to autonomously implement it in their classrooms. Experiments were conducted during schooltime in the classrooms and lasted from February to June 2024.

Preliminary findings indicate good efficiency of the teacher-led training. Secondly, other than the efficiency of the intervention on students' performance, we needed to assess teachers' willingness to adopt such an intervention. At the end of the experiments, interviews were conducted with teachers to evaluate their perception of the acceptability of the intervention and usefulness of the training. In a relevant way, we demonstrated that the workshop effect is robust in ecological setting of classrooms where

the teachers managed themselves the intervention. Indeed, the workshop effects obtained with teachers are similar to those obtained by experimenters in terms of child's curiosity perceptions, metacognitive sensibility and question-asking performances (figure below).



**Figure 37:** Illustration of Workshop effect (Pre vs. Final) on the question asking score when teacher (blue line) or experimenter delivered it.

In related-way, we started in October 2024, experiments to test the relationships the creative thinking-based process and the curious thinking-based processes, with a second field study conducted in three schools within Bordeaux Métropole, involving a total of 130 children aged 9 to 11 years. These experiments aim to further assess the interplay between curiosity and creativity in question-asking behaviors. Drawing from creativity literature, we examine the process of question formulation through associative thinking involved in creativity. To do so, the conversational agent's behavior in "Kids Ask" was modified to prompt children to identify important keywords from a text, then generate free associations based on their prior knowledge. Given the intricate interplay between curiosity and creativity, it is hypothesized that this associative guidance will further enhance children's ability to formulate divergent, curiosity-driven questions.

**Investigating the use of LLM in secondary school.** ChatGPT, one of the most widely used generative AI (LLM) tools, has made accessing mass and personalized information easy and straightforward, even for users without expertise in AI. More particularly, recent reports indicate that the majority of surveyed students aged nine and older have already used this tool for school-related tasks. However, while we know that students are using ChatGPT, there is limited understanding of how they use it and its effects on their learning processes and outcomes, particularly among middle and high school students and in subjects outside programming.

Investigating these patterns of use is a critical step toward identifying the necessary educational interventions to mitigate risks associated with misuse or harmful interactions with ChatGPT, which are particularly likely among non-expert users. To address this, we recruited 63 students aged 14 to 15 and asked them to solve science problems using ChatGPT. We examined their prompt choices, evaluations of ChatGPT's responses, and final problem-solving outcomes. Overall, our results indicate that students are still inefficient users of AI tools such as ChatGPT and are vulnerable to incorporating its misinformation, even when they report high domain knowledge and previous experience with generative AI. This highlights potential misconceptions about these tools' capabilities and the skills required to use them effectively. Furthermore, domain knowledge alone appears insufficient to shield students from adopting misinformation generated by ChatGPT. Implementing formal educational interventions to correct these



misconceptions and train students for informed usage thus seems both timely and essential, given the growing reliance on generative AI tools in education. On the longer term, fostering metacognitive skills may further promote responsible and effective use of such tools (paper in preparation)

### 8.5.3 Supporting Qualitative Analysis with Large Language Models: Combining Codebook with GPT-3 for Deductive Coding

Qualitative analysis of textual contents unpacks rich and valuable information by assigning labels to the data. However, this process is often labor-intensive, particularly when working with large datasets. While recent AI-based tools demonstrate utility, researchers may not have readily available AI resources and expertise, let alone be challenged by the limited generalizability of those task-specific models. In this study, we explored the use of large language models (LLMs) in supporting deductive coding, a major category of qualitative analysis where researchers use pre-determined codebooks to label the data into a fixed set of codes. Instead of training task-specific models, a pre-trained LLM could be used directly for various tasks without fine-tuning through prompt learning. Using a curiosity-driven questions coding task as a case study, we found, by combining GPT-3 with expert-drafted codebooks, our proposed approach achieved fair to substantial agreements with expert-coded results. We lay out challenges and opportunities in using LLMs to support qualitative coding and beyond. This work was published in [190] and involved a collaboration with Z. Xiao, V. Liao, E. Yuan from Microsoft Research Montreal.

## 8.6 Curiosity-driven educational technologies

Since 2019 (Idex cooperation fund between the University of Bordeaux and the University of Waterloo, Canada) and the recent creation of CuriousTECH associate team in 2022 (led by the Flowers team and involving F. Lotte from the Potioc team and M. Fernandes and E. Law from the Waterloo University), we continue our work on the development of new curiosity-driven interaction systems. Substantial progress has been made in this area of application of FLOWERS works (see the website of [CuriousTECH team](#).)

### 8.6.1 New digital approaches for studying curiosity-driven learning

**Participants:** H el ene Sauzeon (*correspondant*), Pierre-Yves Oudeyer (*correspondant*), Rania Abdelghani, Mehdi Alaimi, Fabien Lotte, Aur elien appriou, Myra Fernandes, Edith Law, Yadurshana Sivashankar.

As curiosity is a recent research topic, we studied some basic mechanisms of curiosity-based learning, thanks to three studies have been completed.

The first one regards a new interactive educational application to foster curiosity-driven question-asking in children. Determined to improve children's curiosity, we developed a new interactive system aiming to foster curiosity-related question-asking from texts and their perception of curiosity. To assess its efficiency, we conducted a study with 95 fifth grade students of Bordeaux elementary schools. Two types of interventions were designed, one trying to focus children on the construction of low-level question (i.e. convergent) and one focusing them on high-level questions (i.e. divergent) with the help of prompts or questions starters models. We observed that both interventions increased the number of divergent questions, the question fluency performance, while they did not significantly improve the curiosity perception despite high intrinsic motivation scores they have elicited in children. The curiosity-trait score positively impacted the divergent question score under divergent condition, but not under convergent condition. The overall results supported the efficiency and usefulness of digital applications for fostering children's curiosity that we need to explore further. The overall results are published in CHI'20 [80]. In parallel to these first experimental works, we wrote this year a review of the existing works on the subject [97].

The second study investigates the neurophysiological underpinnings of curiosity and the opportunities of their use for Brain-computer interactions [81]. Understanding the neurophysiological mechanisms underlying curiosity and therefore being able to identify the curiosity level of a person,

would provide useful information for researchers and designers in numerous fields such as neuroscience, psychology, and computer science. A first step to uncovering the neural correlates of curiosity is to collect neurophysiological signals during states of curiosity, in order to develop signal processing and machine learning (ML) tools to recognize the curious states from the non-curious ones. Thus, we ran an experiment in which we used electroencephalography (EEG) to measure the brain activity of participants as they were induced into states of curiosity, using trivia question and answer chains. We used two ML algorithms, i.e. Filter Bank Common Spatial Pattern (FBCSP) coupled with a Linear Discriminant Algorithm (LDA), as well as a Filter Bank Tangent Space Classifier (FBTSC), to classify the curious EEG signals from the non-curious ones. Global results indicate that both algorithms obtained better performances in the 3-to-5s time windows, suggesting an optimal time window length of 4 seconds to go towards curiosity states estimation based on EEG signals. These results have been published [81].

Thanks to Virtual reality device, a third study investigates the role of intrinsic motivation in spatial learning in children [177]. In this study, the state curiosity is manipulated as a preference for a level of uncertainty during the exploration of new virtual environments. To this end, a series of virtual environments have been created and is presented to children. During encoding, participants explore routes in environments according the three levels of uncertainty (low, medium, and high), thanks to a virtual reality headset and controllers and, are later asked to retrace their travelled routes. The exploration area and the wayfinding, i.e. the route overlap between encoding and retrieval phase, (an indicator of spatial memory accuracy) are measured. Neuropsychological tests are also performed. The results showed that there are better performances under the medium uncertainty condition in terms of exploration area and wayfinding score. These first results supports the idea that curiosity states are a learning booster. In Sivashankar et al. study, 10-year-old children (20 females; 22 males) with low to high trait curiosity actively explored virtual environments 39 containing varying levels of uncertainty (low, medium, high) (Fig. 40), after which memory for the route travelled was assessed [177].

As trait curiosity increased (41), so did memory performance in the high uncertainty condition, suggesting that children with high levels of curiosity can better recruit cognitive resources within such environments. Children with high compared to low curiosity also had higher feelings of presence during the immersive experience. Importantly, in environments with medium uncertainty, children with low trait curiosity were able to perform as well as those with high curiosity. Results suggest that individual differences in trait curiosity influences route memory in environments with varying levels of uncertainty.

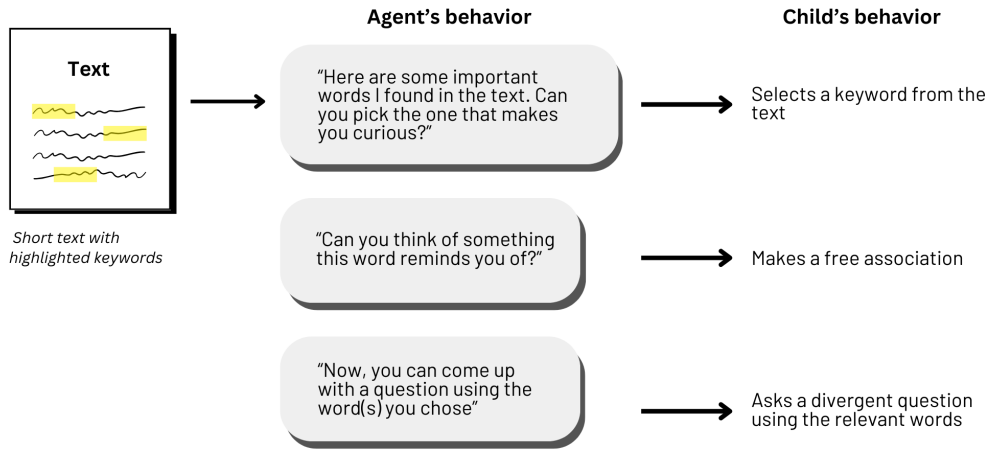
### 8.6.2 Machine Learning for Adaptive Personalization in Intelligent Tutoring Systems

**Participants:** Pierre-Yves Oudeyer (*correspondant*), H el ene Sauzeon (*correspondant*), Benjamin Cl ement, Didier Roy, C ecile Mazon.

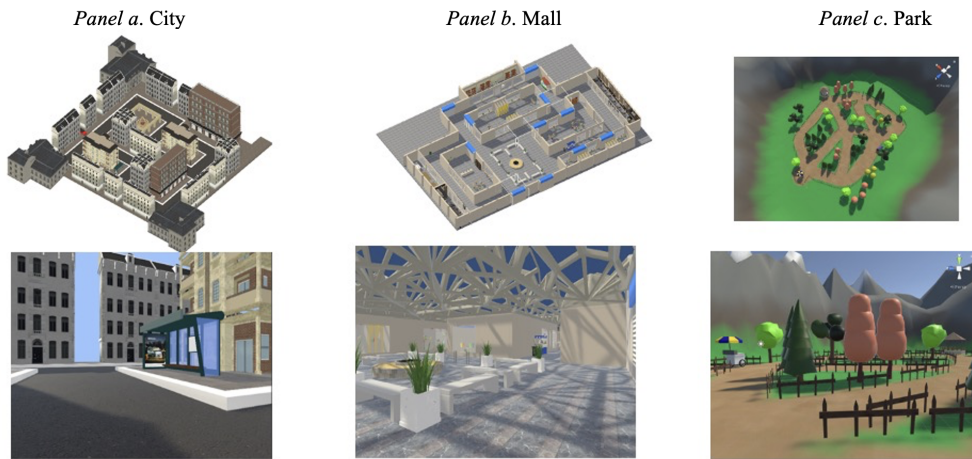
**The Kidlearn project.** is a research project studying how machine learning can be applied to intelligent tutoring systems. It aims at developing methodologies and software which adaptively personalize sequences of learning activities to the particularities of each individual student. Our systems aim at proposing to the student the right activity at the right time, maximizing concurrently his learning progress and his motivation. In addition to contributing to the efficiency of learning and motivation, the approach is also made to reduce the time needed to design ITS systems.

We continued to develop an approach to Intelligent Tutoring Systems which adaptively personalizes sequences of learning activities to maximize skills acquired by students, taking into account the limited time and motivational resources. At a given point in time, the system proposes to the students the activity which makes them progress faster. We introduced two algorithms that rely on the empirical estimation of the learning progress, **RiARiT** that uses information about the difficulty of each exercise and **ZPDES** that uses much less knowledge about the problem.

The system is based on the combination of three approaches. First, it leverages recent models of intrinsically motivated learning by transposing them to active teaching, relying on empirical estimation of learning progress provided by specific activities to particular students. Second, it uses state-of-the-art Multi-Arm Bandit (MAB) techniques to efficiently manage the exploration/exploitation challenge of this optimization process. Third, it leverages expert knowledge to constrain and bootstrap initial



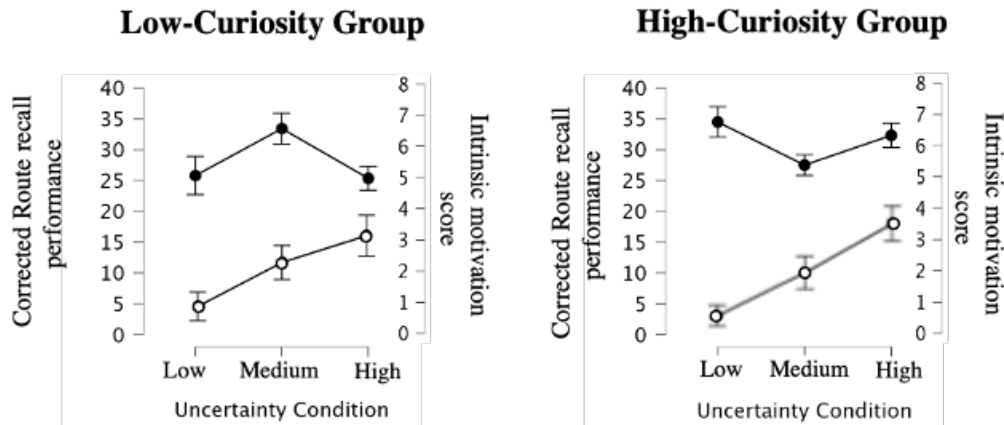
**Figure 38:** Schematic representation of the associative method of prompting in Kids Ask. Children start off by reading a text containing highlighted keywords. They are prompted by the conversational agent to choose one from the list and make a free association with it, based on prior knowledge. They are to use one of or both words to ask a divergent question



**Figure 39:** First-person view and bird's-eye view of the three styles of virtual environments. Participants only experienced the environments from a first-person perspective.



**Figure 40:** From left to right: Condition 1 with Low Uncertainty (1 character); Condition 2 with Medium Uncertainty (3 characters); and Condition 3 with High uncertainty (7 characters)



**Figure 41:** Route Memory Score (black circles) and Intrinsic Motivation Score (white circles) in Low- and High-curiosity Groups as a Function of the Three Uncertainty Conditions (Low, Medium and High)

exploration of the MAB, while requiring only coarse guidance information of the expert and allowing the system to deal with didactic gaps in its knowledge. The system was evaluated in several large-scale experiments relying on a scenario where 7-8 year old schoolchildren learn how to decompose numbers while manipulating money [102]. Systematic experiments were also presented with simulated students.

**Kidlearn Experiments 2018-2019: Evaluating the impact of ZPDES and choice on learning efficiency and motivation.** An experiment was held between March 2018 and July 2019 in order to test the Kidlearn framework in classrooms in Bordeaux Metropole. 600 students from Bordeaux Metropole participated in the experiment. This study had several goals. The first goal was to evaluate the impact of the Kidlearn framework on motivation and learning compared to an Expert Sequence without machine learning. The second goal was to observe the impact of using learning progress to select exercise types within the ZPDES algorithm compared to a random policy. The third goal was to observe the impact of combining ZPDES with the ability to let children make different kinds of choices during the use of the ITS. The last goal was to use the psychological and contextual data measures to see if correlation can be observed between the students psychological state evolution, their profile, their motivation and their learning. We first show that LP-based personalization improves learning performance (reproducing and solidifying previous results) while producing a positive and motivating learning experience. We then show that the addition of self-choice as a playful feature triggers intrinsic motivation in the learner and reinforces the learning effectiveness of the LP-based personalizing. In doing so, it strengthens the links between intrinsic motivation and performance progress during the serious game. Conversely, deleterious effects of the playful feature are observed for hand-designed linear paths. Thus, the intrinsic motivation elicited by a playful feature is beneficial only if the curriculum personalization is effective for the learner. Such a result deserves great attention due to the increased use of playful features in non adaptive educational technologies available in the market. Details of these new results, as well as the overall results of this project, are presented in Benjamin Clément PhD thesis [101] and are currently being processed to be published.

**Kidlearn and Adaptiv'Math.** The algorithms developed during the Kidlearn project and Benjamin Clément thesis [101] are being used in an innovation partnership for the development of a pedagogical assistant based on artificial intelligence intended for teachers and students of cycle 2. The algorithms are being written in typescript for the need of the project. The expertise of the team in creating the pedagogical graph and defining the graph parameters used for the algorithms is also a crucial part of the role of the team for the project. One of the main goal of the team here is to transfer technologies developed in the team in a project with the perspective of industrial scaling and see the impact and the feasibility of such scaling.

**Kidlearn for numeracy skills with individuals with autism spectrum disorders.** Few digital interventions targeting numeracy skills have been evaluated with individuals with autism spectrum disorder (ASD) [145][144]. Yet, some children and adolescents with ASD have learning difficulties and/or a significant academic delay in mathematics. While ITS are successfully developed for typically developed students to personalize learning curriculum and then to foster the motivation-learning coupling, they are not or few proposed today to student with specific needs. The objective of this pilot study is to test the feasibility of a digital intervention using an STI with high school students with ASD and/or intellectual disability. This application (KidLearn) provides calculation training through currency exchange activities, with a dynamic exercise sequence selection algorithm (ZPDES). 24 students with ASD and/or DI enrolled in specialized classrooms were recruited and divided into two groups: 14 students used the KidLearn application, and 10 students received a control application. Pre-post evaluations show that students using KidLearn improved their calculation performance, and had a higher level of motivation at the end of the intervention than the control group. These results encourage the use of an STI with students with specific needs to teach numeracy skills, but need to be replicated on a larger scale. Suggestions for adjusting the interface and teaching method are suggested to improve the impact of the application on students with autism. [142].

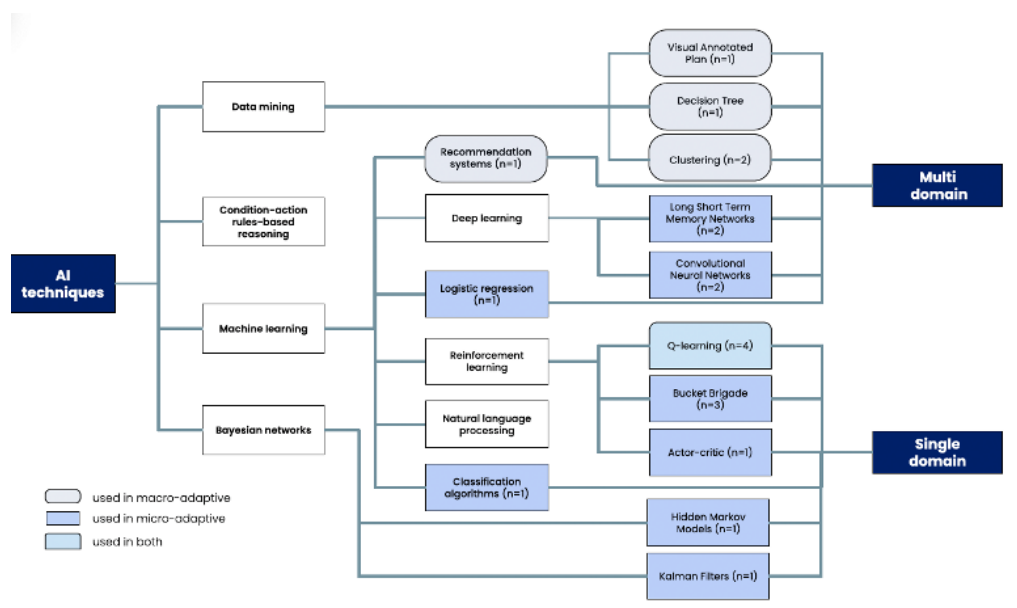
### 8.6.3 Machine learning for adaptive cognitive training

**Participants:** Pierre-Yves Oudeyer, H el ene Sauz eon (*correspondant*), Masataka Sawayama, Benjamin Cl ement, Maxime Adolphe, Marion Pech, Olivier Clerc.

Because of its cross-cutting nature to all cognitive activities such as learning tasks, attention is a hallmark of good cognitive health throughout life and more particularly in the current context of societal crisis of attention. Recent works have shown the great potential of computerized attention training for an example of attention training, with efficient training transfers to other cognitive activities, and this, over a wide spectrum of individuals (children, elderly, individuals with cognitive pathology such as Attention Deficit and Hyperactivity Disorders). Despite this promising result, a major hurdle is challenging: the high inter-individual variability in responding to such interventions. Some individuals are good responders (significant improvement) to the intervention, others respond variably, and finally some respond poorly, not at all, or occasionally. A central limitation of computerized attention training systems is that the training sequences operate in a linear, non-personalized manner: difficulty increases in the same way and along the same dimensions for all subjects. However, different subjects require in principle a progression at a different, personalized pace according to the different dimensions that characterize attentional training exercises.

To tackle the issue of inter-individual variability, the present project proposes to apply some principles from intelligent tutoring systems (ITS) to the field of attention training. In this context, we have already developed automatic curriculum learning algorithms such as those developed in the KidLearn project, which allow to customize the learner's path according to his/her progress and thus optimize his/her learning trajectory while stimulating his/her motivation by the progress made. ITS are widely identified in intervention research as a successful way to address the challenge of personalization, but no studies to date have actually been conducted for attention training. Thus, whether ITS, and in particular personalization algorithms, can optimize the number of respondents to an attention training program remains an open question.

**Grounded state-of-the-art.** To investigate this question, we first conducted a systematic review aiming at exploring existing methods in computerized CT and analyzing their outcomes in terms of learning mechanics (intra-training performance) and effectiveness (near, far and everyday life transfer effects of CT) [86]. A search up to June 2023 with multiple databases selecting 19 computerized CT studies revealed that only two studies emphasized the favorable influence of individualization on CT effectiveness, while five underscored its capacity to enhance the training experience by boosting motivation, engagement, and offering diverse learning pathways. In sum, despite promising results in this new research avenue, more



**Figure 42:** Distribution of AI techniques depending on type of CT studied (multi or single domain) from Adolphe et al., 2024

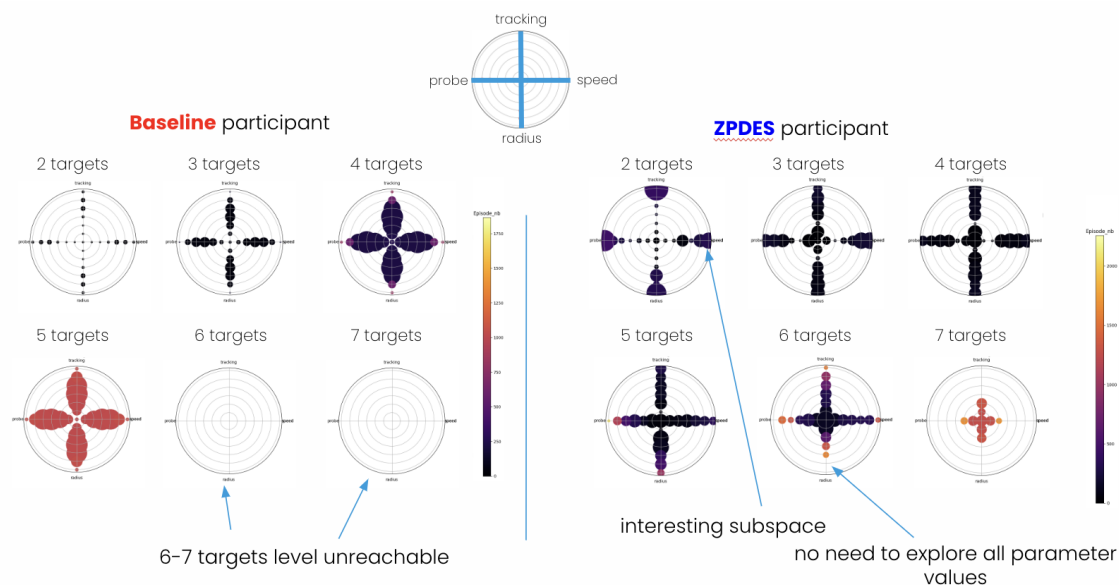
research is needed to fully understand and empirically support individualized techniques in cognitive training.

Complementing the study of adaptive methods applied to cognitive training, we have attempted through a review of the subjective literature to gain a better understanding of the Multiple Object Tracking (MOT) task, which seems to have the best results in terms of attentional training efficiency in young and older adults. Our investigation pursues three main objectives: (1) identifying the cognitive processes influenced by each adjustable parameter of the MOT task; (2) determining which parameters, when progressively adapted during repeated MOT practice, produce the greatest enhancements in task performance; and (3) evaluating how improvements in MOT performance translate into effective transfer effects, including practical, real-world outcomes. The evidence suggests that the MOT task involves a nuanced interplay of visual processing, attentional resources, and working memory, shaped by the intrinsic properties of the objects and the task conditions. The results of this work highlight that: (1) Multiple cognitive mechanisms are identified as active in the task (divided and sustained attention; foveal and peripheral attention; automatic and controlled inhibition, etc.); (2) a limited number of studies have actually implemented the MOT task in computer-assisted cognitive training; and (3) the near (attention tasks) and far (other cognitive tasks) effects that are well documented as positive outcomes of MOT-based training while there is a scarcity of research that has thoroughly analyzed the ecological effects of attentional training, namely the potential transfer effects in everyday life (paper in progress).

**ZPDES calibration for MOT training (Young participants).** In parallel to this, a web platform has been designed for planning and implementing remote behavioural studies. This tool provides means for registering recruited participants remotely and executing complete experimental protocols: from presenting instructions and obtaining informed consents, to administering behavioural tasks and questionnaires, potentially throughout multiple sessions spanning days or weeks. In addition to this platform, a cognitive test battery composed of seven classical behavioural tasks has been developed. This battery aims to evaluate the evolution of the cognitive performance of participants before and after training. Fully open-source, it mainly targets attention and memory. A preliminary study on a large sample of 50 healthy participants showed that the developed tasks reproduced the results of previous studies, that there were large differences between individuals (no ceiling effect) and that the results were significantly reliable between two measurements taken on two days separated by one night [2].

### Randomized and controlled Trial in Young and Olders adults : Predifined vs. ZPDES condition.

Utilizing these tools, a pilot study campaign was conducted to evaluate the impact of our AI-based personalized cognitive training program. The first pilot experiment involved  $n=27$  participants and aimed to compare the effectiveness of a cognitive training program using a linear difficulty management procedure (staircase procedure) to a program using an ITS for difficulty manipulation. The online training lasted for 10 hours over a period of 2 weeks. The results indicated that the ITS-based intervention produced diverse learning trajectories compared to the linear procedure [43](#), leading to broader improvements in pre-post cognitive assessment. However, no significant differences were observed in subjective measures of motivation and engagement between the two groups. Subsequent to this initial experiment, two pilot studies ( $n=11$  and  $n=10$ , respectively) were conducted with the goal of enhancing motivation and engagement in the game. The first study implemented gamified components such as scores and feedback, while the second study examined hyperparameter updates to the ITS. The analysis of learning trajectories, learning outcomes, and subjective measures yielded promising results in favor of the AI-based personalized procedure.



**Figure 43: Different learning trajectories for a selected participant in the staircase group (left) and the ITS group (right). The color of a dot indicates the initial presentation of the parameter value, while the size of the dot represents the frequency of the parameter value.**

Building on the preliminary findings, we expanded our research scope with a more comprehensive experimental setup involving two distinct studies. The first study encompassed 64 young adults, sourced through the Prolific platform, while the second study consisted of 50 older adults, recruited from the "Université du temps libre". Our experimental methodology mirrored that of our initial pilot studies, with a notable enhancement: the integration of new gamified elements (including mini-story creation and new visual content) aimed at boosting participant motivation and engagement.

The data analysis encompassed three primary dimensions: initially, an exploratory phase to delineate learning trajectories between control and intervention groups; subsequently, a comparative analysis of pre- and post-test performance on the cognitive battery; and lastly, an examination of participants' self-reported experiences during training, providing insights into their subjective perceptions of the experiment.

The pilot studies' preliminary outcomes were corroborated in these larger sample groups. Notably, learning trajectories exhibited greater diversity in the group undergoing the intervention procedure. This group also demonstrated a more pronounced improvement across a wider range of cognitive assessment tasks. Although participants engaging in the personalized cognitive training reported a higher cognitive load via questionnaires, the levels of engagement and frustration did not significantly differ between the two groups.

The results showed that ZPDES could be more effective than a control condition, with improved performance on trained tasks in both studies, underlining the benefits of individualized training paths. However, motivation and engagement were lower in the groups using ZPDES, probably due to cognitive load and metacognitive factors. Overall, individualizing cognitive training through systems like ZPDES provides a promising direction for future research by providing automatic methods for taking individual differences into account in CT programs while respecting methodological standards for evaluating the effectiveness of CT. As a result, our work contributes to the growing body of knowledge in both ITS and CT domains while stressing the crucial role of challenges related to motivation and engagement to optimize the effectiveness of these individualized approaches for cognitive and educational outcomes.

As part of the creation of the new University Hospital Institute (UHI) VBHI (VASCULAR BRAIN HEALTH INSTITUTE), we aim to develop and test a personalized, multimodal digital therapeutic approach to slow down the functional consequences of small vessel disease. More specifically:

- Evaluate the impact of personalized cognitive training compared to non-personalized conditions (comparative efficacy).
- Identify potential ElectroEncephaloGraphic (EEG) biomarkers that reflect cognitive activity impacted by small vessel disease and could later (in a subsequent study) be used as targets for exploratory EEG neurofeedback therapy.
- Identify brain areas to target for delivering non-invasive HD-tACS electrical stimulation, using previously acquired MRI data.
- Evaluate the impact of this stimulation on brain activity, neural synchronization, and cognitive performance.

To achieve this, 60 participants from the SHIVA cohort (n=60) will be divided into two subgroups according to the severity of the disease:

- Severe group: presenting multiple lesions on MRI
- Non-severe group: presenting a few lesions

These groups will then be further divided based on the type of training: personalized tests (ZPDES) versus standard tests.

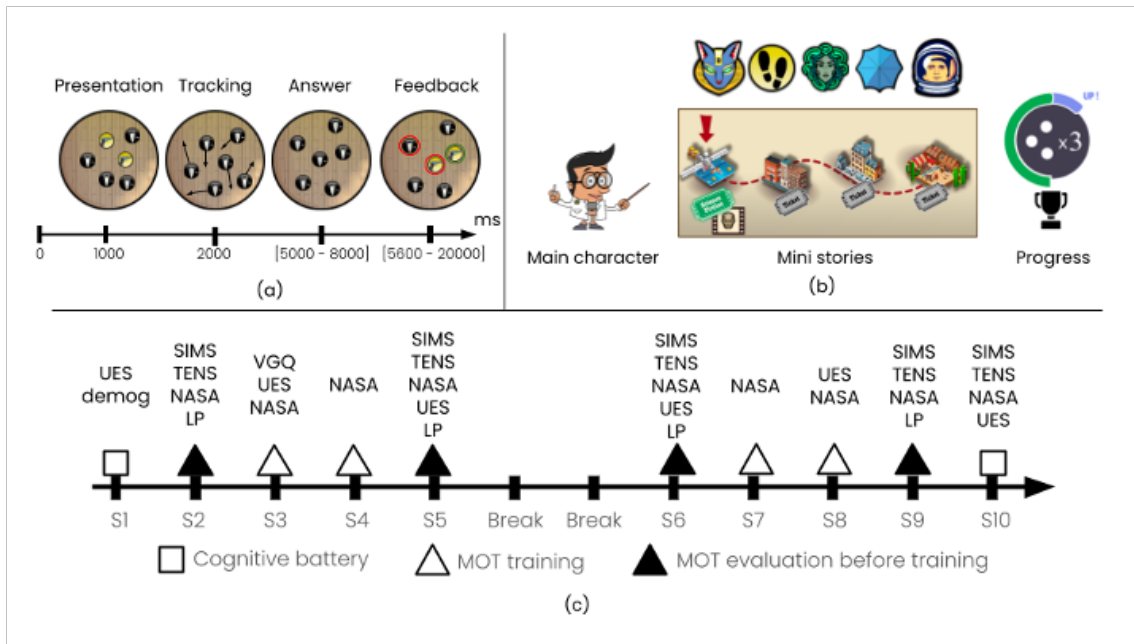
During the pre- and post-training sessions, participants will perform cognitive tests on a computer. Participants will be equipped with an EEG headset, which, combined with a tACS stimulator, will allow for both brain activity recording and stimulation.

We are carrying out an ancillary study with Myra Fernandez's laboratory in Canada, thanks to my participation with the Inria Curiositytech international associate team. We have proposed to collaboratively analyze certain data and dimensions of interest in our respective laboratories (e.g. physical activities) associated with the cognitive training proposed in the SHIVA-DTX-COG project.

**Qualitative Analysis with LLMs:** As it is well known that there are more dropouts in older adults compared to young ones, we aimed to better understand the learning experience of trainees with feedback analyses. For this, we designed a new way through several Large Language Models (LLM) enabling to extract hot topics or main dropout's motivations in verbatim that are related to pragmatic, hedonist and/or aesthetic dimensions of cognitive training. The results analyzed through various LLM are encouraging (paper in progress). To support this new approach, we are exploring different prompts on other data corpora in order to ultimately propose a tutorial accessible to anyone wishing to carry out a LLM-based thematic qualitative analysis.

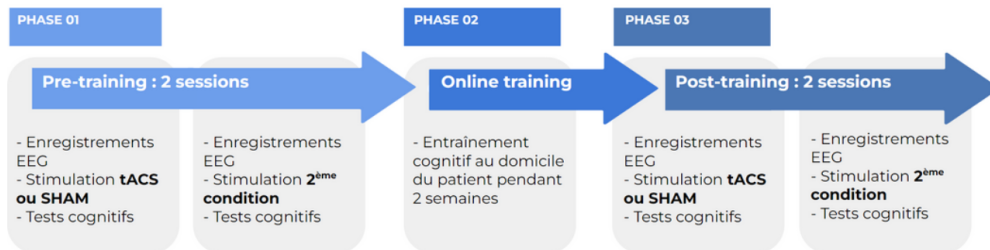
As part of this initiative, a Streamlit application has been created, allowing users to upload their data and perform analyses using LLMs without the need for coding expertise. Additionally, for users with basic coding skills, a series of Jupyter notebooks has been designed. These notebooks leverage the same backend code as the library and provide guided examples to perform various types of qualitative analyses with LLMs, including binary, multiclass, and sequential analyses. Both the application and notebooks are still under development but are already functional.





**Figure 44:** a) The MOT task. (b) Several visual snapshots of intervention. (c) Schedule proposed to participants

Protocole



**Figure 45:** SHIVA study protocol and materials.

#### 8.6.4 ToGather : Interactive website to foster collaboration among stakeholders of school inclusion for pupils with neurodevelopmental disorders

**Participants:** H el ene Sauz eon (*correspondant*), C ecile Mazon, Eric Meyer, Isabeau Saint-Supery, Christelle Maillart (*Uni. Li ege, Belgium*), Kam elia Belassel, Mathieu P eri e, Valentin Strahm.

Sustain and support the follow-up of the school inclusion of children with neurodevelopmental disorders (e.g., autism, attention disorders, intellectual deficiencies) has become an emergency : the higher is the school level, the lower is the amount of schooled pupils with cognitive disabilities.

Technology-based interventions to improve school inclusion of children with neurodevelopmental disorders have mostly been individual centered, focusing on their socio-adaptive, and cognitive impairments and implying they have to adapt themselves in order to fit in our society's expectations. Although this approach centered on the normalization of the person has some advantages (reduction of clinical, symptoms), it carries social stereotypes and misconceptions of cognitive disability that are not respectful of the cognitive diversity and intrinsic motivations of the person, and in particular of the student's wishes in terms of school curriculum to achieve his or her future life project [146].

The "ToGather" project aims at enlightening the field of educational technologies for special education by proposing an approach centered on the educational needs of the students and bringing a concerted and informed answer between all the stakeholders including the student and all their support spheres (family, school, medico-social care). To this end, ToGather project that emanates from participatory design methods, primarily consists of having developed a pragmatic tool (interactive website) to help students with cognitive disability and their caregivers to formalize and to visualize the repertoire of academic skills of the student and to make it evolve according to his or her proximal zone of development (in the sense of Vygotsky) on the one hand, and to the intrinsic motivations of the student (his or her own educational and life project) on the other [143].

This project is in partnership with the School Academy of Bordeaux of the French Education Ministry, the ARI association, the Centre of Autism of Aquitaine. It is funded by the FIRAH (foundation) and the Nouvelle-Aquitaine Region (see the dedicated [webpages](#)).

First, usability studies have been conducted for evaluating ergonomic qualities of the ToGather website, yielding positive results in French and Belgian contexts. Then, we conducted a large field-study to assess the effectiveness of the tool in helping stakeholders to support children with neurodevelopmental disorders (NDD) [169] [167] [168].

The study protocol consisted in a longitudinal non-randomized controlled trial, with baseline, 3-months, and 6-months follow-up assessments. The recruitment was conducted across the entire French territory. Our local partners facilitated the dissemination of the call for participation in Gironde and provided us with contacts to extend it to other regions. Additionally, a recruitment campaign through social media was carried out to communicate about the study and encourage participants to test the ToGather tool.

As the tool was designed to support co-educational process between parents and professionals, a support team had to consist of at least two stakeholders, including at least one of the parents. Initially, 157 participants were recruited in 37 support teams, but 30 individuals did not answer to baseline questionnaire, leading to the exclusion of 11 support teams. After baseline assessment, 13 support teams were allocated to the experimental condition (ToGather app) and 11 to the control condition (usual follow-up).

Primary outcomes measures covered stakeholders' relationships, self-efficacy, and attitudes towards inclusive education, while secondary outcomes measures were related to stakeholders' burden and quality of life, as well as children's school well-being and quality of life.

As the study ended recently, data analysis is still ongoing. Preliminary results after 3 months of use showed encouraging results with an improvement in communication between stakeholders and their respective quality of life (paper in progress)

### 8.6.5 Curious and therefore not overloaded : Study of the links between curiosity and cognitive load in learning mediated by immersive technologies

**Participants:** H el ene Sauz eon (*correspondant*), Matisse Poupard,  
Andr e Tricot (*Cosupervisor - Univ. Montpellier*),  
Florian Larrue (*Industrialist - Le Catie*).

In collaboration with CATIE (industrial partner) and the EPSYLON laboratory at the University of Montpellier (led by Prof. Andr e Tricot), this research program, initiated in April 2022, aims to achieve two main objectives:

- To establish theoretical connections between cognitive load theory and models of curiosity-driven learning.
- To experimentally evaluate how the choice of educational technology influences the relationship between pedagogical approaches (guided instruction vs. exploration) and learner expertise.

To this end, the program is divided into three main phases of study:

**State of the Art.** A systematic review evaluating the contributions and limitations of Virtual Reality (VR) and Augmented Reality (AR) in learning—specifically focusing on their impacts on cognitive load and intrinsic motivation—has been completed and published in the *British Journal of Educational Technology* (BJET) [37].

**Experimentation.** Two experiments were conducted in 2023 with 131 second-year medical students, and replicated in 2024 with 164 medical students ranging from their second to fifth years.

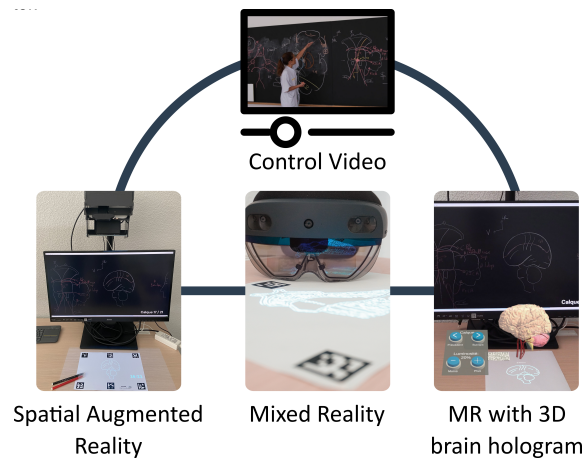
In the first experiment, we used spatial augmented reality (SAR) and mixed reality (MR, HoloLens 2) to investigate whether support students' drawings during lectures could reduce cognitive load and enhance their motivation to learn. Participants followed a 20-minute neuroanatomy video lecture and simultaneously reproduced drawings with the instructor, following four experimental conditions :

- **Spatial Augmented Reality (SAR) Condition:** A digital overlay of the structure to be drawn was projected, allowing the learner to trace the structure onto their paper using a projector and a tracking system.
- **Mixed Reality (MR) Condition:** The digital overlay was projected using a HoloLens2 headset.
- **Mixed Reality with 3D Model (MR+3D) Condition:** In addition to the digital trace, this condition allowed for the display and manipulation of a 3D model of the anatomical structure.
- **Control Condition:** This condition served as the reference and did not include a digital overlay.

Preliminary results from the 2023 data showed that AR-supported drawing led to higher intrinsic motivation, lower extraneous cognitive load, and improved drawing accuracy, particularly in the SAR and MR+3D conditions. However, no significant differences in learning outcomes were observed between conditions. These findings were presented at the 4th International Conference on Science and Technology Education in Porto in October 2024 [75]. The role of students' prior knowledge in anatomy on the effectiveness of different systems was also discussed, and the 2024 data is expected to provide further insights into how expertise influences AR-assisted learning.

In the second experiment, we shifted the learning paradigm by using virtual reality (VR) with varying levels of interaction and guidance. This allowed us to examine how exploration and embodied interaction with a 3D model impact learning, cognitive load, and curiosity.

The analysis of 2023 data, presented in a paper currently under review in *Computers & Education* [69], revealed that VR conditions yielded better learning performance, especially in the passive and active conditions. These were associated with higher intrinsic motivation and optimized cognitive load. Furthermore, intrinsic motivation was positively correlated with germane cognitive load (relevant for

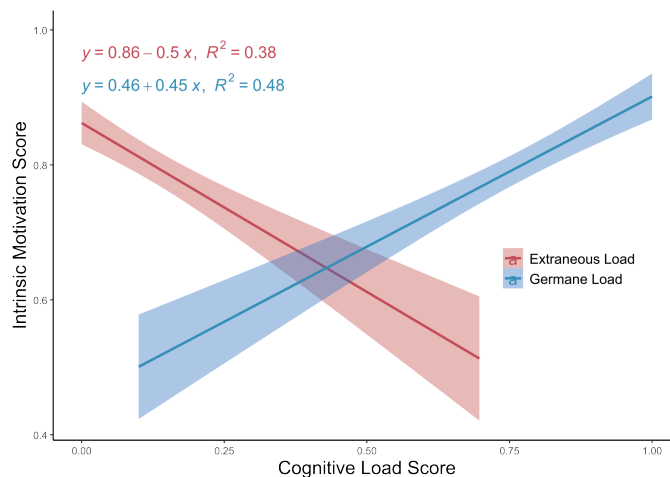


**Figure 46:** Experimental conditions for experiment 1 : Support Drawing with Augmented Reality



**Figure 47:** Experimental conditions for experiment 2 : Embodied learning in virtual reality, effect of interactivity

learning) and negatively correlated with extraneous load (irrelevant for learning). In essence, highly motivated students found the task less cognitively demanding in terms of irrelevant information, enabling them to allocate more cognitive resources to meaningful learning. These findings on the interplay between cognitive load and intrinsic motivation were discussed in a presentation at the International Cognitive Load Theory Conference in Sydney in November 2024 [58].



**Figure 48:** Correlation between intrinsic motivation and cognitive load sub-dimensions

The replication of these two experiments in 2024 aims to deepen our understanding of the relationships between cognitive load and curiosity in learning mediated by immersive technologies, using structural equation modeling."

**Transfer.** We aim to extend the results obtained to the industrial context in which CATIE operates. CATIE's mission is to accelerate technology transfer between the research and industrial sectors.

The Human-Centered Systems team, which supports this research project, assists companies in enhancing the design of existing or new digital systems by applying a human-centered approach. The questions raised by this project are intended to help CATIE address these challenges, refine its expertise in learning and digital systems, and transfer this knowledge to EdTech companies.

#### 8.6.6 Self-determination-driven digital services for supporting aging-in place and well-being: a study of relationships between longitudinal data from smart home and clinical data

**Participants:** H el ene Sauz e on, Juliette Deyts, Lucile Dupuy, Rafik Belloum.

This work is based on data already collected from elderly people with frailty syndrome, who used an ambient assistant-living (HomeAssist) for 12 months, designed to support self-determination and cover 3 domains of need (daily activities, home safety and social participation).

The aim is to analyze various clinical (cognitive assessments, frailty, autonomy, SDT, etc.) and use (user experience questionnaires, diaries, actimetrics based on environmental sensors) data, in order :

- to relate the benefits regarding care with the use of assistive and monitoring services provided by HomeAssist and
- to explore the predictive value of data provided by HomeAssist-based sensors for explaining clinical data

These activities aim to design a user-centered tool that will enable clinicians (psychologists or physicians) to visualize and understand these links effectively. This tool will meet the needs of clinicians

who are currently seeking for tools capable of displaying the evolution of older adults over a long period, in order to detect “weak signals” indicating the first signs of decline. This intuitive visualization will facilitate early intervention, as clinicians will be able to identify trends over time, such as reduced sleep duration, which is a key predictor of health in the elderly.

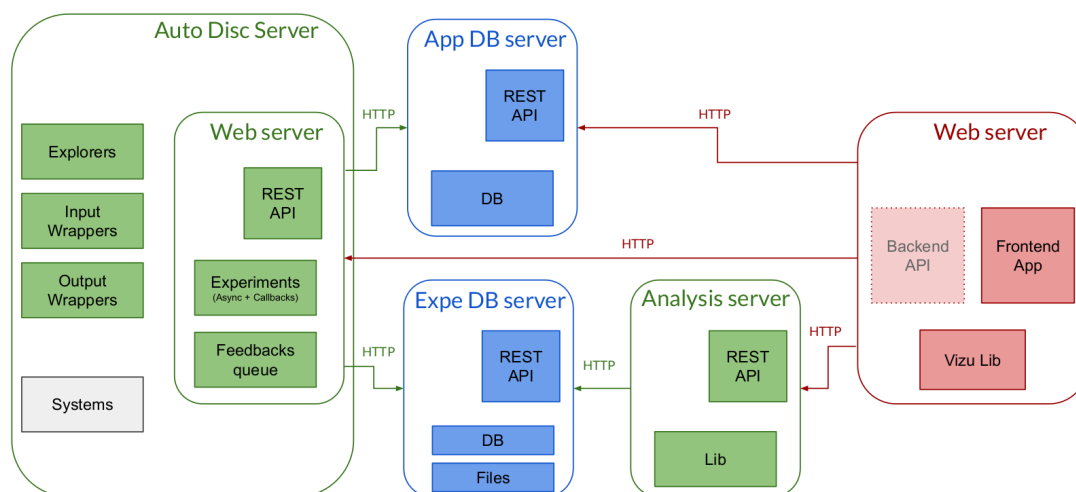
## 8.7 Curiosity-driven AI for assisted scientific discovery

### 8.7.1 Design of an Interactive Software for Automated Discovery in Complex Systems

**Participants:** Clément Romac (*correspondant*), Zacharie Bugaud, Clément Moulin-Frier, Pierre-Yves Oudeyer.

We further developed our Automated Discovery software and particularly focused on adding and experimenting with new systems.

Our **public software** now features more than ten examples ranging from artificial life, to physics or protein docking. The software was publicly released in 2024: [presentation thread](#).



**Figure 49:** Technical architecture of our software.

### 8.7.2 Exploration of Gene Regulatory Network Behaviors Using Automated Discovery Tools

**Participants:** Mayalen Etcheverry (*correspondent*), Clément Moulin-Frier, Pierre-Yves Oudeyer, Michael Levin.

In the context of project "Automated Discovery in Self-Organizing Systems", it has been demonstrated that modern tools leveraging computational models of curiosity developed in the Flowers team can be transposed to form efficient AI-driven "discovery assistants." These tools can assist scientists in mapping and navigating the space of possible outcomes in complex systems [29, 16], [120]. In 2022, we initiated a collaboration with Dr. Michael Levin, a renowned biologist at Tufts University, through a 5-month academic exchange with Mayalen Etcheverry in his lab in Boston. This collaboration laid the foundation for continued collaboration throughout 2023, resulting in the submission of one paper [33] (currently under review) and another accepted at the NeurIPS 2023 *AI for Science* workshop [115].

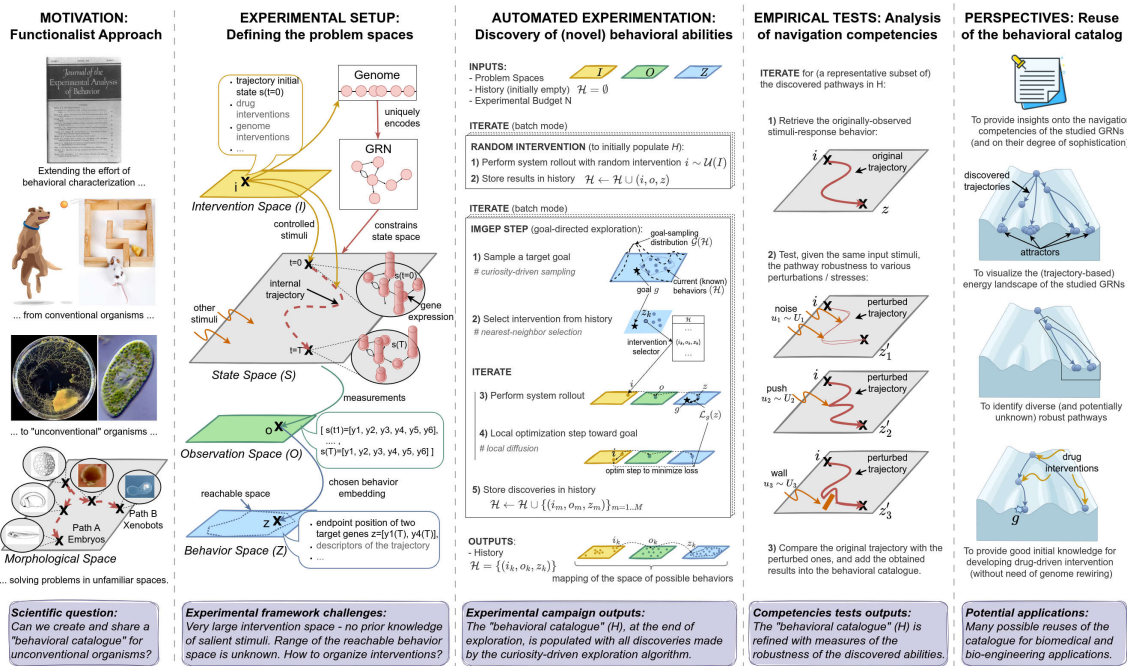


Figure 50: Overview of the proposed framework in [33].

The primary focus of this collaboration was to leverage curiosity-driven exploration algorithms as tools to empower scientific exploration and analysis of basal cognition in biological systems, specifically numerical models of gene regulatory networks (GRNs). Understanding, mapping, predicting, and controlling the complex behavior of these networks is crucial for applications in biomedicine and synthetic bioengineering. However, there are few quantitative tools that facilitate exploration of these networks, especially when their complexity makes unguided exploration infeasible.

To address these challenges in practice, we proposed an experimental framework summarized in Figure 50. In this framework, we formalized and investigated a view of gene regulatory networks as agents navigating a problem space. We developed automated tools to efficiently map the repertoire of robust goal states that GRNs can reach despite perturbations. These tools rely on two main contributions that we made in this work: (1) The use of curiosity-driven exploration algorithms, originating from the AI community, to explore the range of behavioral abilities of a given system, which we adapted and leveraged to automatically discover the range of reachable goal states of GRNs, and (2) The use of a battery of empirical tests inspired by implementation-agnostic behaviorist approaches that we leveraged to assess the navigation competencies of GRNs.

Our data revealed that models inferred from real biological data can reach a surprisingly wide spectrum of steady states, showcasing various competencies that living agents often exhibit in physiological network dynamics and that do not require structural changes to network properties or connectivity. Furthermore, we investigated the applicability of the discovered “behavioral catalogs” for comparing the evolved competencies across classes of evolved biological networks, as well as for the design of drug interventions in biomedical contexts or for the design of synthetic gene networks in bioengineering. Altogether, these automated tools and the resulting emphasis on behavior-shaping and exploitation of innate competencies can open the path to better interrogation platforms for exploring the complex behavior of biological networks in an efficient and cost-effective manner.

To encourage broader adoption and development of the tools and algorithms, we have released two software packages: **SBMLtoODEJax 7.1.17** and **AutoDiscJax**. SBMLtoODEJax converts Systems Biology Markup Language (SBML) models into Python classes written in JAX, enabling easy simulation and manipulation. AutoDiscJax, built upon JAX and SBMLtoODEJax, facilitates automated discovery and exploration of complex systems, specifically organizing the exploration of computational models of biological GRNs.

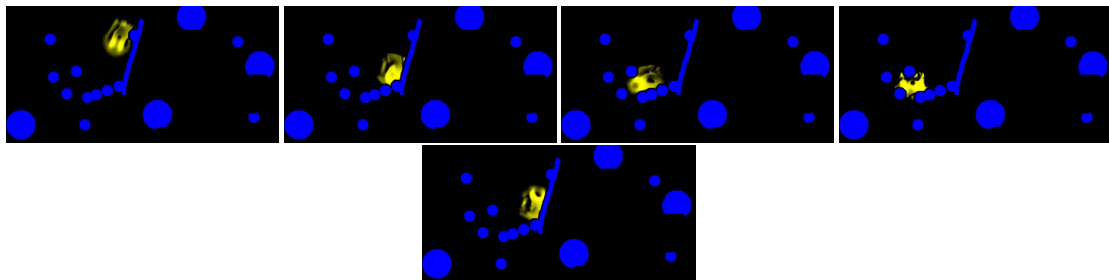
This work has been published in 2024 in the eLife journal [33].

### 8.7.3 Discovering Sensorimotor Agency in Cellular Automata using Diversity Search

**Participants:** Gautier Hamon (*correspondant*), Mayalen Etcheverry, Bert Chan, Clément Moulin-Frier, Pierre-Yves Oudeyer.

As a continuation of the previous projects in Automated Discovery in Self-Organizing Systems, we have been working on expanding the set of discoveries of possible structures in continuous CAs such as Lenia [99, 98], and in particular we have been interested to search for emerging *agents* with *sensorimotor* capabilities. Understanding what has led to the emergence of life and sensorimotor agency as we observe in living organisms is a fundamental question. In our work, we initially only assume environments made of low-level elements of matter (called atoms, molecules or cells) locally interacting via physics-like rules. There is no predefined notion of agent embodiment and yet we aim to answer the following scientific question: *is it possible to find environments in which there exists/emerge a subpart that could be called a 'sensorimotor agent'?*

We use Lenia continuous cellular automaton as our artificial "world" [98]. We introduce a novel method based on gradient descent and curriculum learning combined within an intrinsically-motivated goal exploration process (IMGEP) to automatically search parameters of the CA rule that can self-organize spatially localized<sup>1</sup> and moving patterns<sup>2</sup> within Lenia. The IMGEP defines an outer exploratory loop (generation of training goal/loss) and an inner optimization loop (goal-conditioned). We use a population-based version of IMGEP [17],[106] but introduce two novel elements compared to previous papers in the IMGEP literature. First, whereas previous work in [29] and [16] used a very basic nearest-neighbor goal-achievement strategy, our work relies on gradient descent for the local optimization of the (sensitive) parameters of the complex system, which has shown to be very powerful. To do so we made a differentiable version of the Lenia framework, which is also a contribution of this work. Secondly, we propose to control subparts of the environmental dynamics with functional constraints (through predefined channels and kernels in Lenia) to build a curriculum of tasks; and to integrate this stochasticity in the inner optimization loop. This has shown central to train the system to emerge sensorimotor agents that are robust to stochastic perturbations in the environment. In particular, we focus on modeling obstacles in the environment physics and propose to probe the agent sensorimotor capability as its performance to move forward under a variety of obstacle configurations. We also provide in this work tests and metrics to measure the robustness of the obtained agents.



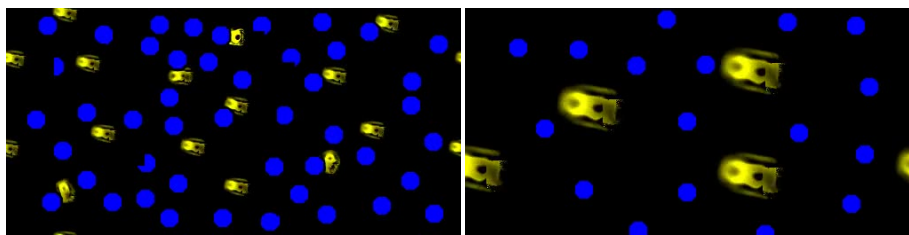
**Figure 51:** Robustness test to harder/unseen obstacle configurations: straight wall, bigger obstacle, dead ends.

While many complex behaviors have already been observed in Lenia, among which some could qualify as sensorimotor behaviors, they have so far been discovered "by chance" as the result of time-consuming manual search or with simple evolutionary algorithms. Our method provides a more systematic way to automatically learn the CA rules leading to the emergence of basic sensorimotor structures, as shown in Figure 53. Moreover, we investigated and provided ways to measure the

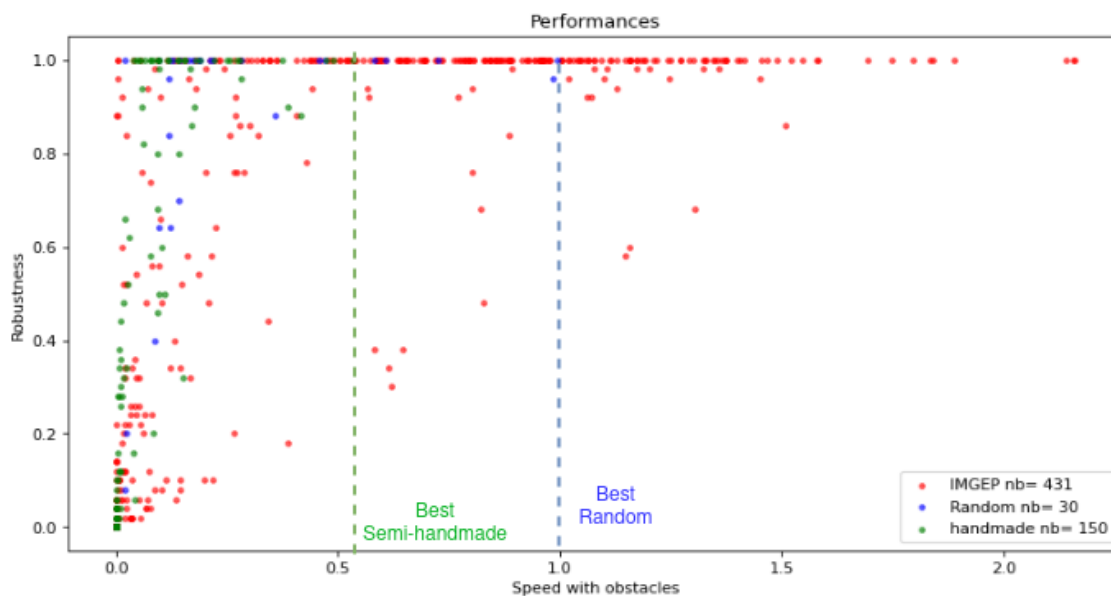
<sup>1</sup>**Spatially-localized pattern** a pattern existing within some (fuzzy) boundary i.e. with a limited range in space as opposed to patterns with unbounded growth

<sup>2</sup>**Moving patterns** a spatially-localized pattern that move and propagate information in space





**Figure 52:** Change of scale changing the kernel size and initialization, the grid is the same size in both

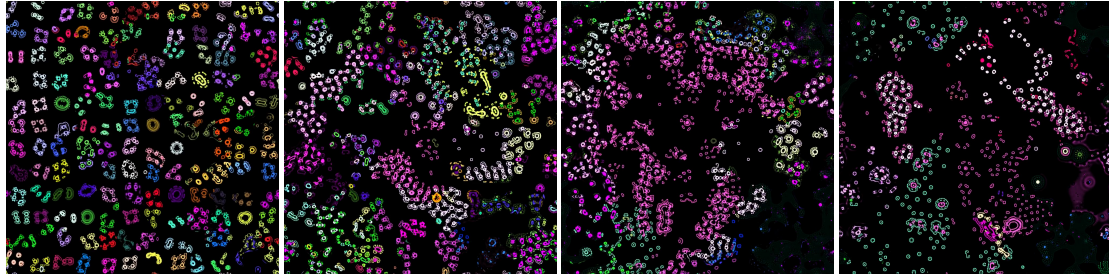


**Figure 53:** Scatter plot of the agents as their measured performances of robustness to obstacles (y axis) and speed in obstacles (x axis) obtained by IMGEP (red), random search with the same compute resources as IMGEP (blue) and the one from the original lenia paper (green)

(zero-shot) generalization of the discovered sensorimotor agents to several out-of-distribution perturbations that were not encountered during training. Impressively, even though the agents still fail to preserve their integrity in certain configurations, they show very strong robustness to most of the tested variations. The agents are able to navigate in unseen and harder environmental configurations while self-maintaining their individuality (Figure 51). Not only the agents are able to recover their individuality when subjected to external perturbations but also when subjected to internal perturbations: they resist variations of the morphogenetic processes such that less frequent cell updates, quite drastic changes of scales as well as changes of initialization (Figure 52). Furthermore, when tested in a multi-entity initialization and despite having been trained alone, not only the agents are able to preserve their individuality but they show forms of coordinated interactions (attractiveness and reproduction). Our results suggest that, contrary to the (still predominant) mechanistic view on embodiment, biologically-inspired embodiment could pave the way toward agents with strong coherence and generalization to out-of-distribution changes, mimicking the remarkable robustness of living systems to maintain specific functions despite environmental and body perturbations [131]. Searching for rules at the cell-level in order to give rise to higher-level cognitive processes at the level of the organism and at the level of the group of organisms opens many exciting opportunities to the development of embodied approaches in AI in general.

The work has been released in 2022 as a distill-like article which is currently hosted [at this link](#). This article contains an interactive demo in WebGL and javascript, as well as many videos and animations of the results. A colab notebook with the source code of the work is publicly available [at](#).

In 2024, additional quantitative experiments were conducted as well as ablations. This work has been



**Figure 54:** Multi-species simulation in Flow Lenia where each colour represents different values of parameters. Left to right shows the evolution of the system over time with some species stealing the mass of others.

submitted to the Science Advances journal.

#### 8.7.4 Flow lenia: Mass conservation for the study of virtual creatures in continuous cellular automata

**Participants:** Erwan Plantec, Gautier Hamon (*correspondant*), Mayalen Etcheverry, Pierre-Yves Oudeyer, Clément Moulin-Frier, Bert Chan.

Following our work on trying to find sensorimotor capabilities in cellular automata such as Lenia [99, 98], we kept exploring the search for low level cognition in continuous cellular automata. This led to preliminary search on trying to emerge memory in self-organizing agents as well as work on trying to implement other environmental constraints in the CA in order to emerge interesting behavior. To implement more easily those environmental constraints as well as to ease the emergence of spatially localized patterns (and thus have the optimization/search to focus more on the cognitive ability, removing the need to optimize to prevent uncontrollable growth/explosion of the pattern), we worked on adding mass conservation to the Lenia system.

We propose in this work a mass-conservative (i.e the sum of the CA's activations remains constant over time) extension to Lenia called Flow Lenia [164]. We hypothesize that such conservation laws will help in the search for artificial life-forms by constraining emerging patterns to spatially localized ones. It also allows to implement more easily environmental constraints on the self-organizing agents such as a need for food to grow, etc.

Furthermore, we show that this new model allows for the integration of the update rule parameters within the CA dynamics enabling the emergence of creatures with different parameters and so different properties in the same environment/grid. This leads to multi-species simulation where the grid is filled with agents with different behaviors and properties 54. Such a feature opens up research perspectives towards the achievement of open-ended intrinsic evolution inside continuous CAs, which means that all the evolutionary part would be a result of the dynamic of the CA (without any external loop/system). We hypothesize that this open-ended intrinsic evolution could, through the competition/cooperation, lead to the emergence of interesting low level cognition in those system.

Simple evolutionary strategy (with an evolutionary loop outside the system) was also used to optimized for pattern with directional and rotational movement.

You can find some examples of the system and pattern in this companion website, including the ones trained for movement, random parameters, food in flow Lenia, and multi species s,imulations: [see](#). Notebook with the system can be found : [here](#).

This work led to an oral presentation to the WIVACE 2022, 15th International Workshop on Artificial Life and Evolutionary Computation.

In 2023, final quantitative experiments on optimizing the parameters with evolutionary strategies and writing was conducted, as well as some additional exploratory experiments on large simulations for open ended evolution.

This work got the best paper award at the ALIFE 2023 conference where it got an oral presentation.

In 2024, an extended version of the paper was accepted in the *Alife Journal* (submitted upon their demand), with additional experiments focusing on the emergent evolutionary dynamics of the simulation.

**Collaboration with Bert Chan.** In the context of the project Automated Discovery in Self-Organizing Systems, we have an ongoing collaboration with Bert Chan, a previously independent researcher on Artificial Life and author of the Lenia system [99, 98] and who is now working as a research engineer at Google Brain. During this collaboration, Bert Chan help us design versions of IMGEP usable by scientists (non ML-experts) end-users, which is the aim of project 8.7.1. Having himself created the Lenia system, he is highly-interested to use our algorithms to automatically explore the space of possible emerging structures and provides us valuable insights into end-user habits and concerns. Bert Chan also co-supervised with Mayalen Etcheverry the master internship of Gautier Hamon which led to the work described in section 8.7.3. He also co-supervised with Gautier Hamon and Mayalen Etcheverry the master internship of Erwan plantec which led to the work described in section 8.7.4.

### 8.7.5 Semantic Open-Endedness in Flow-Lenia using Vision Language Models and IMGEP

**Participants:** Sina Khajehabdollahi (*correspondent*), Gautier Hamon, Cédric Colas, Pierre-Yves Oudeyer, Clément Moulin-Frier.

The growing capabilities of large language models (and their extension into vision language models) has opened the possibility to integrate these foundation models within the framework of automated discovery. Using Flow-Lenia as the substrate for a complex and emergent system, we integrate vision language models into an IMGEP exploration algorithm in order to discover behavioural patterns in goal spaces that are semantically meaningful and interpretable. Previous work on exploration without these foundation models required expert knowledge and human intervention in order to hand-craft the goal spaces and/or targets for exploration and optimization. In cases where goal spaces were not hand-crafted, the semantic interpretation of automatically clustered behaviours remained difficult to translate to human perception. By utilizing vision language models to label patterns generated by FlowLenia, the output behaviour of the model can be summarized using keywords or sentences which are then readily interpretable by humans. Setting behavioural goals can then also be done at a high level with language, where we use these higher-level abstractions to push for more biologically rich behaviours that would otherwise be difficult or impossible to quantify in a generalized way. In this framework, we can search for open-endedness in these models of complexity in a much broader and human-relevant way, leveraging human expectations of novelty and higher-level descriptors through the eyes of these emerging foundation models.

## 9 Bilateral contracts and grants with industry

### 9.1 Bilateral contracts with industry

#### Research on lifelong Deep Reinforcement Learning of multiple tasks (Microsoft)

**Participants:** Pierre-Yves Oudeyer (*correspondant*), Laetitia Teodorescu.

Financing of the PhD grant of Laetitia Teodorescu.

#### Automated Discovery of Self-Organized Structures (Poïetis)

**Participants:** Pierre-Yves Oudeyer (*correspondant*), Mayalen Etcheverry.

Financing of the CIFRE PhD grant of Mayalen Etcheverry by Poietis.

### Machine learning for adaptive cognitive training (OnePoint)

**Participants:** H el ene Sauz eon (*correspondant*), Pierre-Yves Oudeyer, Maxime Adolphe.

Financing of the CIFRE PhD grant of Maxime Adolphe by Onepoint.

### Curiosity-driven interaction system for learning (evidenceB)

**Participants:** H el ene Sauz eon (*correspondant*), Pierre-Yves Oudeyer, Rania Abdelghani.

Financing of the CIFRE PhD grant of Rania Abdelghani by EvidenceB.

### Curious and therefore not overloaded : Study of the links between curiosity and cognitive load in learning mediated by immersive technologies (CATIE)

**Participants:** H el ene Sauz eon (*correspondant*), Matisse Poupard, Andr e Tricot (*Cosupervisor - Univ. Montpellier*), Florian Larrue (*Industrialist - Le Catie*).

Financing of a PhD grant of Matisse Poupard with CATIE and EPSYLON Lab (Univ. Montpellier).

### Augmenting curiosity-driven exploration with very large language models in deep reinforcement learning agents (Hugging Face)

**Participants:** Pierre-Yves Oudeyer (*correspondant*), Cl ement Romac.

Financing of the PhD grant of Cl ement Romac by Hugging Face.

### Autonomous Driving Commuter Car (Renault)

**Participants:** David Filliat (*correspondant*), Emmanuel Battesti.

We developed planning algorithms for a autonomous electric car for Renault SAS in the continuation of the previous ADCC project. We improved our planning algorithm in order to go toward navigation on open roads, in particular with the ability to reach higher speed than previously possible, deal with more road intersection case (roundabouts), and with multiple lane roads (overtake, insertion...).

We received a 30keuros grant from Google Brain, as well as 30keuros Google cloud credit, for developing projects on automated exploration of continuous cellular automata.

## 9.2 Bilateral Grants with Foundation

### School+ /ToGather project (FIRAH and Region Nouvelle-Aquitaine)

**Participants:** H el ene Sauz eon (*correspondant*), C ecile Mazon, Isabeau Saint-supery, Eric Meyer.

Financing of one year-postdoctoral position and the app. development by the International Foundation for Applied Research on Disability (FIRAH). The School+ project consists of a set of educational technologies to promote inclusion for children with Autism Spectrum Disorder (ASD). School+ primary aims at encouraging the acquisition of socio-adaptive behaviours at school while promoting self-determination (intrinsic motivation), and has been created according to the methods of the User-Centered Design (UCD). Requested by the stakeholders (child, parent, teachers, and clinicians) of school inclusion, Flowers team works to the adding of an interactive tool for a collaborative and shared monitoring of school inclusion of each child with ASD. This new app will be assessed in terms of user experience (usability and elicited intrinsic motivation), self-efficacy of each stakeholder and educational benefit for child. This project includes the Academie de Bordeaux –Nouvelle Aquitaine, the CRA (Health Center for ASD in Aquitania), and the ARI association.

#### **CLEMENCE Cohort (Fondation de France and Théa Pharma)**

**Participants:** Hélène Sauzéon (*correspondant*), Cécile Mazon, Cécile Delcourt.

The project "Cohorte Longitudinale sur la Myopie et le développement oculaire dans l'ENfanCE(CLEMENCE) is led by C. Delcourt from the lab of Bordeaux Populational Health (2M€). Hélène Sauzéon and Cécile Mazon participate to the research program with the study of developmental changes due to Myopia in visual attention.

## **10 Partnerships and cooperations**

### **10.1 International initiatives**

#### **10.1.1 Associate Teams in the framework of an Inria International Lab or in the framework of an Inria International Program**

##### **CuriousTECH Inria Associate team**

**Participants:** Helene Sauzéon, co-PI, Inria, Edith Law, Co-PI,UW, Pierre-Yves Oudeyer, Inria, Cecile Mazon , Inria, Marion Pèch , Inria, Matisse Poupard, Inria, Maxime Adolphe, Inria, Myra Fernandes, UW.

<https://flowers.inria.fr/curioustech-associate-team>

**Title:** CURIOS TECH

**Partner Institution(s):**

- HCI Lab, University of Waterloo, Canada
- Cognitive Neuroscience Lab, University of Waterloo, Canada

**Date/Duration:** 2023/ 03/1 to 2025/12/31

**Additional info/keywords:** CURIOSTech team aims to develop an original, cross-disciplinary approach, joining together two perspectives: Objective 1) The fundamental study of curiosity-driven learning across the life-span (in children, young adults and older adults) ; and Objective 2) The study of how new (re)educational technologies, using both curiosity-related models and artificial intelligence techniques, can personalize learning sequences for each individual, maximizing curiosity and learning efficiency in real world contexts.

Marion Pech, Maxime Adolphe and Matisse Poupard have visited our UW colleagues for developing our joint works (from december 2023 to January 2024). Similarly during the last december, we wellcame two UW student, i.e., Sophia Tran (PhD Student, Cog. Neuroscience Lab) and Amelia Mewdsley (PhD student, Cog. Neuroscience Lab)

Marion Pech. December 11, 2024 visits from Nathalie Bier, professor of occupational therapy at Université de Montréal, and Patricia Belchior, professor of occupational therapy at McGill University. These researchers work in the field of normal and pathological cognitive aging and cognitive intervention, notably using new technologies. Exhibitions on the work of Equipe Flowers, Projet ZPDES and thesis research at INSERM. Brainstorming for collaborations.

## 10.2 International research visitors

### 10.2.1 Visits of international scientists

**Gaia Molinaro**, PhD student at University of Berkeley, visited the team for 6 months (Feb-July), working on several projects including: developing the concept of latent learning progress to account for human curiosity-driven exploration, and testing it in a novel experimental paradigm with humans (this led to a publications at Neurips 2024, [50]), collaborating on a project about autotelic generative models (publications at Neurips 2024 [38]), and on a project modeling the dynamics of iterated transmission of stories in chains of LLMs (publications at ICLR 2025, [67]). This visit was partially funded by a grant from the Chateaubriand fellowship program. We also welcomed **Anne Collins**, associate professor at University of Berkeley.

Max Taylor-Davis, PhD student at the University of Edimburg, made a 3-month visit during the summer 2024. This led to the publication of a paper accepted at the EVOSTAR 2025 conference (see New Results).

We also welcomed several researchers for short visits and invited seminars: Yves Barral (ETH Zurich, feb. 2024), Edith Law (Univ. Waterloo, Canada, feb. 2024), Mathieu Lefort (Univ. Lyon, May 2024), Louise Goupil (CNRS, Grenoble, France).

### 10.2.2 Visits to international teams

#### Research stays abroad

##### **Gautier Hamon**

**Visited institution:** Univeritat Pompeu Fabra (UPF), Barcelona, Complex Systems Lab

**Country:** Spain

**Dates:** 6 January 2024 to 6 April 2024

**Context of the visit:** Project on "The emergence of agriculture in artificial agent populations" in collaboration with Ricard Solé. (Benefitted from the UBGRS-Mob grant from the Collège des Ecoles Doctoral de l'Université Bordeaux)

**Mobility program/type of mobility:** research stay

## 10.3 European initiatives

### 10.3.1 Horizon Europe

**INTERACT** [INTERACT project on cordis.europa.eu](https://cordis.europa.eu/project/interact)

**Title:** Help Me Grow: Artificial Cognitive Development via Human-Agent Interactions Supported by New Interactive, Intrinsically Motivated Program Synthesis Methods.

**Duration:** From October 1, 2022 to August 31, 2026

#### Partners:

- INSTITUT NATIONAL DE RECHERCHE EN INFORMATIQUE ET AUTOMATIQUE (INRIA), France
- MASSACHUSETTS INSTITUTE OF TECHNOLOGY (MIT), United States

**Inria contact:** Cedric Colas

**Summary:** Building machines that interact with their world, discover interesting interactions and learn open-ended repertoires of skills is a long-standing goal in AI. This project aims at tackling the limits of current AI systems by building on three families of methods: Bayesian program induction, intrinsically motivated learning and human-machine linguistic interactions. It targets three objectives: 1) building autonomous agents that learn to generate programs to solve problems with occasional human guidance; 2) studying linguistic interactions between humans and machines via web-based experiments (e.g. properties of human guidance, its impact on learning, human subjective evaluations); and 3) scaling the approach to the generation of constructions in Minecraft, guided by real players. The researcher will collaborate with scientific pioneers and experts in the key fields and methods supporting the project. This includes supervisors Joshua Tenenbaum (program synthesis, MIT) and Pierre-Yves Oudeyer (autonomous learning, Inria); diverse collaborators, and an advisory board composed of an entrepreneur and leading scientists in developmental psychology and human-robot interactions. The 3rd objective will be pursued via a secondment with Thomas Wolf (CSO) at HuggingFace, a world-leading company in the open source development of natural language processing methods and their transfer to the industry. By enabling users to participate in the training of artificial agents, the project aims to open research avenues for more interpretable, performant and adaptive AI systems. This will result in scientific (e.g. interactive program synthesis approaches), societal (e.g. democratized AI training) and economic impacts (e.g. adaptive AI assistants). The dissemination, communication and exploitation plans support these objectives by targeting scientific (AI, cognitive science), industrial (video games, smart homes) and larger communities (gamers, software engineers, large public).

### 10.3.2 Other european programs/initiatives

ORA project - Open Research Area (ORA) for the Social Sciences 8th call for proposals

**Title:** How curiosity enhances learning across childhood and adolescence: The role of metacognition and agency

**Duration:** From January 5, 2024 to December 31, 2027

**Partners:**

- INSTITUT NATIONAL DE RECHERCHE EN INFORMATIQUE ET AUTOMATIQUE (INRIA), France
- Cardiff University, UK
- MaxPlanck institute, Berlin, Germany

**Inria contact:** Pierre-Yves Oudeyer

**Coordinator:** Mathias Grüber, Brain and Imagery centre, Cardiff University, UK ; Funds : 1,177 k€

**Summary:** 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. Initial research suggests that curiosity not only enhances learning but also metacognitive awareness of one's own learning progress strengthens curiosity. However, the reciprocal influences between curiosity-based learning and metacognition, particularly during late childhood and adolescence when both abilities continue to develop, remain poorly understood. To address this gap in the existing literature, our interdisciplinary approach aims to explore the bidirectional connections between curiosity-based learning and metacognition during this crucial developmental period. Additionally, late childhood and adolescence is marked by increasing needs of agency, which can boost the curiosity effects on learning because the learner is in control of the learning material. We will therefore seek to study how the protracted development of metacognition enhances the efficiency of curiosity-based learning, especially in situations of high agency. To this end, we will conduct a series of five experiments, leveraging complementary

skills and approaches of world-leading research teams around curiosity and development. We will combine different methodological approaches, including behavioral experiments and functional neuroimaging, training studies, and longitudinal assessments to investigate how across-person differences and within-person changes in metacognition and agency contribute to curiosity-based learning across late childhood and adolescence. We will leverage established and recently developed experimental and naturalistic paradigms from our labs to understand how metacognition affects curiosity-based learning across development. Furthermore, we will translate the lab-based findings from the proposed experiments to design pedagogical interventions that stimulate curiosity and metacognition in the classroom. This is a timely and innovative project as recent research provides all the building blocks needed for a step change in our understanding of the mechanisms of curiosity development. The complementary perspectives of the three pioneering teams (Cardiff, Bordeaux, Trier) provide a unique collaborative opportunity to combine parallel research domains to generate important discoveries on the development of curiosity-based learning with broad scientific and societal impact.

Clément Moulin-Frier is continuing his collaboration with Ricard Solé (Universitat Pompeu Fabra, ICREA, Spain) and Marti Sanchez-Fibla (CSIC, Spain) on using Multi-Agent Reinforcement Learning to study fundamental questions on the origins of eco-engineering strategies in the human species. In 2024, this collaboration led to the 3-months visit of Gautier Hamon (Flowers PhD student) in the Complex Systems Lab of Ricard Solé, with a paper in preparation on the emergence of agricultural practices in RL agents. It also led to the publication of a journal article in the *Journal of the Royal Society Interface* on the cooperative control of environmental extremes [39].

#### 10.4 National initiatives

**ANR Chaire Individuelle Deep Curiosity** - PY Oudeyer continued to work on the research program of this Chaire, funding 2 PhDs and 3 postdocs for five years (until 2025).

**ANR JCJC ECOCURL** - C. Moulin-Frier obtained an ANR JCJC grant. The project is entitled "ECOCURL: Emergent communication through curiosity-driven multi-agent reinforcement learning". The project starts in Feb 2021 for a duration of 48 months. It will fund a PhD student (36 months) and a Research Engineer (18 months) as well as 4 Master internships (one per year).

**Projet AixIA: "Analyse d'Interférences par Intelligence Artificielle"**. Pierre-Yves Oudeyer and Clément Moulin-Frier obtained a grant from the call for project AIRSTRIP "L'intelligence Artificielle au service de l'Ingénierie des Systèmes Aéronautiques et spatiaux", in collaboration with the IRT Saint Exupéry. The project was accepted in 2023 and will fund 18 months of a research engineer position starting in 2024.

**Inria Exploratory Action AIDE** - Didier Roy is collaborator of the Inria Exploratory Action AIDE "Artificial Intelligence Devoted to Education", ported by Frédéric Alexandre (Inria Mnemosyne Project-Team), Margarida Romero (LINE Lab) and Thierry Viéville (Inria Mnemosyne Project-Team, LINE Lab). The aim of this Exploratory Action consists to explore to what extent approaches or methods from cognitive neuroscience, linked to machine learning and knowledge representation, could help to better formalize human learning as studied in educational sciences. AIDE is a four year project started middle 2020 until 2024 *see*.

**Inria Exploratory Action l'AM** - Hélène Sauzéron is co-PI with P. Dragicevic of the Inria Exploratory Action l'AM "Impact of Augmented Reality on Autobiographical Memory: Examining Involuntary Memories and False Memories" (174,5k€). Starting in last september, the aim of this Exploratory Action consists to explore to what extent augmented reality based devices can produce erroneous autobiographical memories, and more particularly in vulnerable people (Children and older adults or young adults with low memory abilities of source monitoring).



**New collaboration with Maxime Derex from IAST Toulouse** for the co-direction of the PhD thesis of Jeremy Perez with Clément Moulin-Frier and Pierre-Yves Oudeyer on "Interactions between intrinsically motivated goal-exploration processes and cumulative cultural evolution" (see section 8.2.5).

**France 2030 - PPR AUTONOMIE : Vieillesse Et Situations De Handicap - Projet INNOVCare (Lechevalier S., 3,5M€) (2023-26)** - H  l  ne Sauz  on and AS Rigaud will supervise the WP5 dedicated to two care-led innovation experiments with assistive technologies (400k   for Bordeaux). - H  l  ne Sauz  on is responsible of the WP3 « Digital technology for aging in place » (470k /3,5M ), D  fi 4 - Num  rique, Innovcare (PPR Autonomie PIA2030, 2023-28).

**VBHI project(Vascular Brain Health Institute -IHU, led by S. Debette, 5M ) (2023-26)** - H  l  ne Sauz  on will supervise the WP4.3 dedicated to "Explore Digital Therapeutics To Slow Down Cognitive Decline In Covert Csvd" (150k )

#### 10.4.1 Adaptiv'Math

- Adaptiv'Math
- Program: PIA
- Duration: 2019 - 2020
- Coordinator: EvidenceB
- Partners:
  - EvidenceB
  - Nathan
  - APMEP
  - LIP6
  - INRIA
  - ISOGRAD
  - Daesign
  - Schoolab
  - BlueFrog

The solution Adaptiv'Math comes from an innovation partnership for the development of a pedagogical assistant based on artificial intelligence. This partnership is realized in the context of a call for projects from the Ministry of Education to develop a pedagogical platform to propose and manage mathematical activities intended for teachers and students of cycle 2. The role of Flowers team is to work on the AI of the proposed solution to personalize the pedagogical content to each student. This contribution is based on the work done during the Kidlearn Project and the thesis of Benjamin Clement [101], in which algorithms have been developed to manage and personalize sequence of pedagogical activities. One of the main goal of the team here is to transfer technologies developed in the team in a project with the perspective of industrial scaling.

## 11 Dissemination

### 11.1 Scientific events: organisation

#### Member of the organizing committees

- C  dric Colas, Laetitia Teodorescu and Gaia Molinaro co-organized the **Intrinsically Motivated Open-ended Learning workshop** at Neurips 2024.

- Hélène Sauzeon was a member of the organization committees of 3 scientific events:
  1. **International Summer School of Disability Alliance between FEDRHA & REPAR** (Canadian Researcher Network for Disability) about « **Handicap, mobility and XR: Beyond the real?** », May 27-30th, 2024, Marseille (*Institut du Mouvement, Université Aix-Marseille*);
  2. **Annual Conference of FEDHRA**, November 7th & 8th, 2024, at the Centre of Inria of the University of Bordeaux, Talence (*Journées scientifiques de la FEDRHA: “Le handicap mental à tous les âges de la vie: vivre avec et l’accompagner”*);
  3. **2nd Annual Conference of Gérontopôle Nouvelle Aquitaine**, April 16th, 2024, *Cité internationale de la BD et de l’image*, Angoulême.
- Cécile Mazon was a member of the organization committees of 2 scientific events
  1. **Annual Conference of FEDHRA**, November 7th & 8th, 2024, at the Centre of Inria of the University of Bordeaux, Talence (*Journées scientifiques de la FEDRHA: “Le handicap mental à tous les âges de la vie: vivre avec et l’accompagner”*);
  2. **Inria Scientific days on Disability**, November 19th & 20th, 2024, at the Hôtel-Dieu, Paris.
- PY Oudeyer co-organized the Life, Structure and Cognition symposium 2024 at IHES, France, this year on the topic "A Multitude of Times", [link](#).
- PY Oudeyer and H. Sauzéon co-organized the Workshop on Children learning at school through curiosity and creativity at Inria Bordeaux.

#### Member of the conference program committees

- Hélène Sauzéon was a member of the conference program committees of the **International Conference on Advances in Computer-Human Interactions** (since 2022).
- Hélène Sauzéon attended the program committees of 2 scientific events:
  1. **Annual Conference of FEDHRA**, November 7th & 8th, 2024, at the Centre of Inria of the University of Bordeaux, Talence (*Journées scientifiques de la FEDRHA: “Le handicap mental à tous les âges de la vie: vivre avec et l’accompagner”*);
  2. **2nd Annual Conference of Gérontopôle Nouvelle Aquitaine**, April 16th, 2024, *Cité internationale de la BD et de l’image*, Angoulême.
- PY Oudeyer was senior area chair of the Reinforcement Learning Conference 2024.

**Reviewer** Clément Romac has reviewed for ICML workshops, EWRL, NeurIPS conference and workshops. Jérémy Perez has reviewed for Evolutionary Human Sciences. Grgur Kovač has reviewed for Artificial Intelligence Review. Nicolas Yax has reviewed for PLOS Computational Biology and Nature Communications Psychology. Sina Khajehabdollahi has reviewed for the ICLR 2024 conference. Matisse Poupard has reviewed for Metaverse journal and NeurIPS workshop. Helene Sauzéon reviewed 3 conference papers for International Conference on Advances in Computer-Human Interactions and ACM Interaction Design & Children. Cécile Mazon reviewed articles for Nature Scientific Reports, Education and Information Technology, and International Journal of Developmental Disabilities

**Member of the editorial boards** PY Oudeyer was member of the editorial board of: Frontiers in Neurorobotics.

### 11.1.1 Invited talks

Hélène Sauzéon gave 2 invited talks:

1. **Les technologies numériques au service d'un vieillissement en bonne santé**, 13ième Édition **Conférence Handicap (IFRATH)**, 5–7 juin, Paris;
2. CuriousTECH associate team: From the study of curiosity-driven learning across life-span to the design of new interactive curiosity-driven systems for learning in real-life settings. **Workshop Inria–University of Waterloo–Université de Bordeaux**, 22–23rd February, Bordeaux, France.

PY Oudeyer gave 9 invited talks:

- (mars 2024) **Curiosité et métacognition: sciences cognitives, IA et éducation**, Groupe de travail sur la métacognition du Conseil Scientifique de l'Education Nationale.
- (mars 2024) **Grounding Language and LLMs through Autotelic Curiosity-driven Learning**, Seminar in Cognitive Informatics, DIC, UQÀM, Canada.
- (June 2024) Les enjeux de la recherche en IA pour l'éducation, Conseil Scientifique de l'Education Nationale
- (June 2024) Open-ended learning and curiosity in infants and machines: learning progress, autotelism and language, Workshop on Emotions and Learning in infants, University Paris Nanterre
- (July 2024) Generating a Diversity of Challenging Programming Puzzles with Autotelic Generative Models, Workshop LLM4Code, Inria Paris.
- (Oct 2024) The learning progress hypothesis: a theory of curiosity-driven learning and applications in educational technologies, Workshop on Leveraging educational games to study motivation, University of Amsterdam.
- (Nov 2024) Open-ended generation of diverse solvable coding problems with autotelic generative models, **Workshop on Program Synthesis**, Labri, University of Bordeaux.
- (Nov 2024) Les enjeux sociétaux de l'éducation à l'IA générative, Journée "Enfance et numérique: pour un numérique éducatif au service de la réussite de tous les élèves", associée à la chaire Neuroéducation et créativité (G. Borst), Sorbonne Université.
- (Jan 2025) Curiosity-driven learning in humans: learning progress, autotelic exploration and open-ended development, Keynote lecture at the **Budapest Conference on Cognitive Development**.

### 11.1.2 Leadership within the scientific community

Hélène Sauzéon was member of scientific committee of **Gérontopôle Nouvelle-Aquitaine**, since 2021

Helene Sauzéon supported the structuration of inria research network on Disability and the valorization research outcomes for the internal road map for Inria as socio-organization since 2023

Helene sauzéon is member of GIS Institut de la Longévité, des Vieillessees et du Vieillissement (ILVV), since 2021 and of **RTP CNRS Éducation** network since 2022.

Hélène Sauzéon and PY Oudeyer are members of two local NA networks related to digital research and its applications (R3IA et R3NumEd) since 2022

Cécile Mazon is co-responsible of the WP on technological tool in the PIA AtypieFriendly

PY Oudeyer collaborated with J. Gottlieb, R. Wilson and E. Bonawitz on the design of new conference series in the Gordon Research Conference program, dedicated to the Sciences of Curiosity.

PY Oudeyer continued to be part of the steering committee of the Life, Structure and Cognition conference series at IHES, which is a yearly interdisciplinary symposium organized at IHES, and where every year he participates to define the main scientific topics and to select and invite speakers from various disciplines.

PY Oudeyer was part of the advisory committee of the **Intrinsically Motivated Open-ended Learning workshop** at Neurips 2024.

PY Oudeyer led the design of the AI dimensions of the research and development program of major P2IA (ministry of education) and BPI projects in educational technologies, in collaboration with the EvidenceB company.

### 11.1.3 Scientific expertise

Hélène Sauzéon was member of HCERES Evaluation Committee of the ChArT Lab (Paris 8) in 2024.

Hélène Sauzéon was member of international committee of "**Fonds Rech. Québec**" for assessing the Canadian multidisciplinary research network on ageing, 2024

Hélène Sauzéon was a reviewer for the ANRT (2 PhD projects)

Hélène Sauzéon is member of the ANR Committee for "interface" Topic (CE 38 pluridisciplinary committee: Computer, Humanities, and social sciences) since 2023

Hélène Sauzéon participated to a working group for instructing the creation of an Inria Project Team (new-STARS, for the center of Nice -Sophia Antipollis) in 2024

Hélène Sauzéon is scientific expert of the Research Project Call by **Fondation Clément Fayat** (since 2022).

Clément Moulin-Frier reviewed a project for the ANR.

PY Oudeyer reviewed projects for the Templeton Foundation and the European Commission, and was a reviewer of the CIFAR **research program on Learning in Brains and Machines** (Nov. 2024).

### 11.1.4 Research administration

PY Oudeyer has been head of the Flowers project team, and coordinated the design and writing of a new research program for the renewal of the project-team at Inria.

Hélène Sauzéon is local referent at the Inria Centre of UB for **PIQ** since 2024

Hélène Sauzéon is responsible of the Research axis "**Interventions innovantes**" in ACTIVE team of the BPH Lab (UB-Inserm) since 2022

Hélène Sauzéon is the head of the Innovations and Transfer Committee of the **BIND Center of Excellence** in Bordeaux and member of the BIND directory Committee since 2018

Hélène Sauzéon is member of Scientific Committee of extended committee of Project-team of the centre of Inria of university of Bordeaux, since 2020.

Hélène Sauzéon is member of directory committee of IFHR which is a national institute on disability funded by Inserm aiming the researcher networking and dissemination of knowledge on multidisciplinary research on disability; since 2018.

Helene Sauzéon was member of scientific committee of **MAVIE-II-Calyxis**.

## 11.2 Teaching - Supervision - Juries

### 11.2.1 Teaching

Founded Teaching Innovation projects

- Hélène Sauzéon is responsible of a task within the i UB2030-CAP santé numérique - France 2030 -AMI 2022 Compétences et Métiers d'avenir, Volet santé numérique (€3.1m) (Coord., R. Thibault). This will involve applying and testing algorithms for personalizing learning paths on certain e-learning courses in health and digital technology for the Bordeaux University (€120k).
- Hélène Sauzéon is member of the "Jeu, Inclusivité & Hanpathie (JIH)" project (Coord., E. Dugas, in charge of Mission Handicap, Bordeaux University) - AAP Région NA 2023- "Favoriser la réussite étudiante" (Promoting student success). The aim of this educational project is to design a digital game to develop empathy towards students with disabilities and thus contribute to a better quality of social life for these students (interactions with their peers and University staff).

Teaching Responsibilities in University structures:

- Clément Moulin-Frier is responsible professor of the "System Design, Integration and Control" course at the University Pompeu Fabra in Barcelona, Spain.
- Cécile Mazon is responsible of the second year of the master curriculum in Technology, Ergonomics, Cognition and Handicap (TECH, Cognitive Sciences - University of Bordeaux) since sept. 2021.
- Cécile Mazon is responsible of the master curriculum in Technology, Ergonomics, Cognition and Handicap (TECH, Cognitive Sciences - University of Bordeaux) since sept. 2022.
- Cécile Mazon is responsible of the licence curriculum in Cognitive sciences (Licence MIASHS) since sept. 2024.

Teaching Involvement in Computer / Engineer science or in cognitive science for Bachelors and master Degrees:

- BS & Master: Cognitive Science, Univ. of Bordeaux- , 104,5h, Marion Pech
- BS & Master: Neuropsychology, Univ. of Caen-,4H, Marion Pech
- Licence MIASHS,Univ. of Bordeaux: Supervision for research, (6h) Marion Pech
- ENSC/ENSEIRB Presentation of developmental artificial intelligence and the Flowers Lab, 2h, Option Robot (Laetitia Teodorescu)
- ENSC Introduction to bayesian analysis, 8h, Option AI (Adolphe Maxime)
- ENSC Transformers and Large Language Models, 16h, Option AI (Clément Romac)
- ENSC Reinforcement Learning, 8h, Option AI (Clément Romac)
- BS & Master: Cognitive Science, Univ. of Bordeaux- , 240h, Cécile Mazon
- Master: Navigation for Robotics, 21 h, M2, ENSTA Paris, David Filliat
- Master: Navigation for Robotics, 24 h, M2 DataAI, IP Paris - Paris, David Filliat
- Université de Bordeaux - MIASHS Bachelor : Models and measures of high-level cognitive functions - knowledge and representation, 22h, (Poupard Matisse)
- Université de Bordeaux - TECH Master: Virtual reality, interaction and healthcare applications, 7h, (Poupard Matisse)
- Université de Bordeaux - TECH Master: Scientific basis, 2h, (Poupard Matisse)
- 2nd year : Deep Learning, 12h, IMT Atlantique (Sao Mai Nguyen).
- Université de Bordeaux - TECH Master: IT Project Management, 18h (Isabeau Saint-Supery)
- Université de Bordeaux - TECH Master: IT Project Management, 14h (Marie-Sarah Desvaux)
- Master UPF-Barcelona: Robotics and AI, 10h (Clément Moulin-Frier)
- PY Oudeyer gave a course on developmental reinforcement learning at ENSEIRB (2h), dec. 2024.
- PY Oudeyer gave a course on developmental reinforcement learning at Cogmaster at Sorbonne Univ./ENS (2h+4h project jury), dec. 2024.
- Master: Cognitive Science, 24h (Eric Meyer)
- Teacher training at the Rectorat of Bordeaux : digital technologies for students with special educational needs , 6h (Eric Meyer)
- 2nd year Master in cognitive science : Assistive technologies (20h), Rania Abdelghani.

- University Degree in Neuropsychological Sciences : Cognitive aging and Digital technologies (3h), Helene Sauz on
- Cl ment Moulin-Frier and Corentin L ger gave a course on System Design, Integration and Control at the University Pompeu Fabra in Barcelona, Spain (7.5h). This year, we used for the first time the new Vivarium simulator (see New Software).

Teaching Involvement in Computer / Engineer science or in cognitive science for lifelong curriculum

- H l ne Sauz on and Marie-Sarah Desvaux gave talks and animated workshops for ' cole Acad mique de la Formation Continue' of Rectorat of Paris, on the subject of Curiosity-driven learning in schools
- H l ne Sauz on and Matisse Poupard gave a talk and animated a workshops on "User-centered design of XR applications" for ' cole d' t  2024 - Mobilit  et r alit  virtuelle : au del  du r el' (FEDRHA/REPAR)
- University Degree in Neuropsychological Sciences : Cognitive aging and Digital technologies (3h), Helene Sauz on
- University Degree in Neuropsychological Sciences : Neuropsychological assessment in normal aging (3h), C cile Mazon

### 11.2.2 Supervision

- **Rania ABDELGHANI** (50%, codirection avec P.Y. Oudeyer, DR Inria) Guider les esprits de Demain : Agents Conversationnels pour Entra ner la Curiosit  et la M tacognition chez les Jeunes Apprenants, depuis Octobre 2020 (CIFRE avec EvidenceB), **defended**, Sept. 27th, 2024.. (supervisors: H. Sauz on & P.Y. Oudeyer)
- **Maxime ADOLPHE** (50%, codirection avec P.Y. Oudeyer, DR Inria) Design and evaluation of new personalisation algorithms embedded in an intelligent tutorial system for attention training / Conception et  valuation de nouveaux algorithmes de personnalisation embarqu s dans un syst me tutoriel intelligent d'entra nement de l'attention" beg. in sept. 2020 (CIFRE contract with Onepoint), **defended**, Sept. 27th, 2024. (supervisors: H. Sauz on & P.Y. Oudeyer)
- PhD in progress: Grgur Kovac, "Developmental training of socio-cognitive abilities in AI systems", (supervisors:PF Dominey and P.Y. Oudeyer)
- PhD in progress: Gauthier Hamon, "Open-endedness in artificial life and artificial intelligence: an eco-evo-devo perspective" (supervisor: C. Moulin-Frier)
- PhD in progress: Nicolas Yax, "Studying cognitive and metacognitive skills in foundation models" (supervisors: S. Palminteri, P.Y. Oudeyer)
- PhD in progress: Julien Pourcel, "Autotelic LLMs that learn how to code", (supervisors: C. Moulin-Frier and P.Y. Oudeyer)
- PhD in progress: Thomas Carta, "LLM-based Autotelic deep reinforcement learning agents", (supervisors: O. Sigaud, S. Lamprier and P.Y. Oudeyer)
- PhD in progress: Cl ment Romac, "Grounding LLMs with online RL", (supervisors: T. Wolf and P.Y. Oudeyer)
- PhD in progress: Jeremy Perez, "Studying mechanisms and roles of curiosity in socio-cultural contexts" (supervisors: C. Moulin-Frier, M. Derex, P.Y. Oudeyer)
- PhD in progress: Timoth  Boulet, "Controller synthesis for artificial agents in simulated environments using generative AI" (supervisors C. Moulin-Frier, X. Hinault, N. Fijalkow)

- PhD in progress: Marko Cvjetko, "Autotelic exploration algorithms for automated search of open-endedness in artificial life" (supervisors: C. Moulin-Frier, PY. Oudeyer) item PhD in progress: Loris Gaven, "Metacognitive prediction of learning progress for guiding autotelic agents" (supervisors: PY. Oudeyer and C. Moulin-Frier)
- PhD in progress: **Isabeau SAINT-SUPERY** "Designing and Assessing a new interactive tool fostering stakeholders' cooperation for school inclusion", beg. in sept. 2021, supervised by H. Sauz on and C. Mazon.
- PhD in progress : **Matisse POUPARD** "Optimize learning in a digital environment according to learners' level of expertise, epistemic curiosity and mode of instruction" ", beg. in April. 2022 (CIFRE Contract with Catie), supervised by H. Sauz on and A. Tricot (Univ. Montpellier).
- PhD in progress : **L ana PETIOT** (50% codirection avec P. Dragicevic, CR Inria) Impact of augmented reality on involuntary autobiographical memory /Impact de la r alit e augment e sur la m moire autobiographique involontaire, beg. in. Oct. 2023 (Allocation Inria AEx IAM, 2023-25), supervised by H. Sauz on and P. Dragicevic (Potioc Team, CR Inria)
- PhD in progress : **Chlo e (Marie-Sarah) DESVAUX** (100%) Designing and testing digital metacognitive training workshops to encourage curiosity and creativity in a school context Concevoir et exp rimer des ateliers num riques d'entra nement m tacognitifs pour favoriser la curiosit  et la cr ativit  en contexte scolaire, Beg. Oct. 2023 (Alloc. MESRI, ED SP2 -Univ. Bordeaux), supervised by H. Sauz on and PY Oudeyer
- PhD in progress : Juliette DEYTS "HomeAssist : study of actimetric and clinical data from 131 equipped participants" (Alloc. Sur Projet ANR Innovcare, ED SP2 -Univ. Bordeaux), d but Nov. 2024, supervised by H. Sauz on and Lucile Dupuy, MCF (BPH lab, Active team)

### 11.2.3 Juries

- PY Oudeyer was a member of the selection committee of the Inria Prize from Acad mie des Sciences.
- PY Oudeyer was reviewer in the PhD of Jerome Tinker (OIST, Japan), James Daly (Univ. Texas, US) and Jean-Baptiste Gaya (Sorbonne Universit ).
- PY Oudeyer was in the PhD "comit  de suivi" of Yanis Bendi-Ouis (Univ. Bordeaux), Paul-Antoine Le Tolguennec (Univ. Toulouse), Matthis Poupard (Univ. Bordeaux), Marie Martin (Universit  Interdisciplinaire de Paris), Marc Welter (Univ. Bordeaux),
- Cl ment Moulin-Frier was reviewer of the PhD thesis of Timoth e Anne (thesis director: Jean-Baptiste Mouret) at the Universit  of Lorraine (France), defended on June 6, 2024.
- Cl ment Moulin-Frier was reviewer of the PhD thesis of Maxime Toquebiau (thesis director: Nicolas Br d che) at Sorbonne Universit  (ISIR lab), defended on December 13, 2024.
- Cl ment Moulin-Frier was examiner of the PhD thesis of Lo c Goasguen (thesis director: Sylvain Argentieri) at Sorbonne Universit  (ISIR lab), defended on December 18, 2024.
- Cl ment Moulin-Frier was examiner of the PhD thesis of Noami Chaix-Eichel (thesis director: Nicolas Rougier) at the Universit  of Bordeaux, defended on November 22, 2024.
- H. Sauz on is member of 6 PhD juries (2 president, 3 reviewer ; 1 examiner).
  - Jianling Zou Laboratoire Cognitions humaine et artificielle (EA4004) / Uni. Paris 8 President
  - Christelle Nahas Equipe TheMAS - laboratoire TIMC (UMR 5525) / Uni. Grenoble Reviewer
  - Hugo Fournier Equipe Sant  - LaB ; de Psychologie / UB President
  - Elen Sargzyan Equipe ELIPSE – Lab. IRIT / University Paul Sabatier Toulouse Reviewer
  - Edwige Chauvergne Equipe Bivouac – Centre Inria de l'UB Examiner

– Julie Restout UR 20217 HAVAE -Institut Recherche Omegahealth / Univ. Limoges Reviewer

- H. Sauzéon and C. Mazon are permanent members of jury of Master degree in cognitive science at the university of Bordeaux.

### 11.3 Popularization

Dialog with policy makers

H. Sauzéon and PY. Oudeyer were interviewed and wrote reports to contribute to the [report of French Senate on AI and education](#).

PY Oudeyer participated to several events organized with/by the educational institutions (Ministry of education) to train teachers, and more broadly civil servants, to artificial intelligence (mechanisms, applications and societal dimensions):

- Presentation on "[IA gé'ne'rative, socie'te' et e'ducation: En quoi l'IA gé'ne'rative repre'sente-elle un enjeu dans la formation des citoyens ?](#)", for the [Conférence de Consensus on Nouveaux Savoirs et Nouvelles Compétences des Jeunes](#) of Cnesco, associated to the [written note](#) on the same topic (Nov. 2024)
- Introductory course on artificial intelligence to a group of high-school teachers from Nouvelle-Aquitaine (Inria, Jan. 2024)
- Presentations of the state of scientific research in training curiosity and metacognition in children using AI models (Académie de Nouvelle Aquitaine, Cardie, journées IA et éducation, March and April, 2024)
- Course on "[Usages de l'IA en éducation: retours d'expériences et perspectives](#)", MOOC AI4T
- Participation to a debate on [Comment augmenter la confiance en soi des élèves](#) for an audience of middle and high-school teachers, July 2024.
- [Course on generative artificial intelligence](#) (history, mechanisms, applications, societal dimensions) for a group of senior officials, members of cabinets of various ministries, organized by DINUM
- Presentation on AI and education for the training day "Etapp-IA" dedicated to head of middle and high-schools and "inspecteurs d'académie", Académie de Nouvelle Aquitaine (Jan. 25)

Clément Romac gave a talk at the French Ministry of Economics and Finance on "AI in 2024" for the LaborIA.

Dialog with civil society

- Clément Romac and Jérémy Perez gave talks on AI during a training day for teachers (November 2024).
- Matisse Poupard led a "BUREAU DES ENQUETES" workshop at the Cap Sciences" museum, in which visitors had to guess his thesis topic from three objects.
- Eric Meyer, Cécile Mazon and Hélène Sauzéon gave a talk about together app for supporting school inclusion ; «Journée d'étude - Outils numériques et besoins éducatifs particuliers at INSEI in April, 24th 2024 (event for parents, teachers)
- Hélène Sauzéon gave a talk about the educational technologies for curiosity for the Council of teachers in secondary and high schools" (collège des IA-IPR) of Nouvelle Aquitaine Academia, July, 5th 2024, Lycée des Graves à Gradignan
- Hélène Sauzéon gave a popularized talk for 2 classes of Year 11 at the Lycee Saint Louis in the context of CHICHE -1 classe-1scientifique
- All Flowers members took part in welcoming the 3rd year trainees (secondary school).



- H  l  ne Sauz  on and Marie-Sarah Desvaux gave talks and animated workshops for '  cole Acad  mique de la Formation Continue' of Paris, on the subject of Curiosity-driven learning in schools
- H  l  ne Sauz  on and Matisse Poupard gave a talk and animated a workshop on "User-centered design of XR applications" for '  cole d'  t   2024 - Mobilit   et r  alit   virtuelle : au del   du r  el' (FEDRHA)
- Cl  ment Moulin-Frier was invited at the M  diath  que de Blanquefort on March 21, 2024, to give a talk and led a large-audience discussion in the context of the Mois de la Connaissance. The talk was entitled : Qu'est-ce-que l'intelligence (artificielle) ?
- PY Oudeyer participated to the Chiche program, visiting several classrooms of the high-school Kaestler (Mars 2024)
- PY Oudeyer welcomed and presented the team's work to several groups of middle school intern students.

### 11.3.1 Specific official responsibilities in science outreach structures

PY Oudeyer participated several times to the piloting committee of the research network on education and digital sciences (R3NumEd, Nouvelle Aquitaine)

### 11.3.2 Productions (articles, videos, podcasts, serious games, ...)

D. Roy and P-Y. Oudeyer wrote a popular science book to introduce generative AI (mechanisms, applications, societal dimensions) to adolescents, as well as to their teachers and families. It is entitled "C'est (pas) moi, c'est l'IA", and was published in September 2024 by Nathan. It was reviewed in widely distributed magazines (e.g. Magazine de l'APEL) and radios (e.g. France Culture, RFI). The web page of the book is here: <http://www.pyoudeyer.com/cli/>.

A. Torres-Legu  t, C. Romac, T. Carta and PY. Oudeyer produced the pedagogical video series "ChatGPT explained in 5 mn", aimed at training generative AI literacy in a wide diversity of students (e.g. high school), available here: [https://developmentalsystems.org/chatgpt\\_en\\_5\\_minutes/](https://developmentalsystems.org/chatgpt_en_5_minutes/). They are under a Creative Commons licence, CC-BY, enabling open and free reuse. They were already integrated in the MOOC AI4T (<https://www.ai4t.eu>), as well as in an internal training platform of "Acad  mie du Num  rique du Minist  re de la d  fense", in a mobile app made by Inria with educational materials related to AI (<https://epoc.inria.fr/epocs/E009LL/>), and are being adapted and integrated in a training platform for the whole population of civil servants in France, coordinated by DINUM.

Cl  ment Romac, Thomas Carta and Pierre-Yves Oudeyer wrote down the article "Les algorithmes des IA peuvent-ils comprendre notre monde ?" published in Pour La Science n  557 (February 2024).

Cl  ment Moulin-Frier, Gautier Hamon and Pierre-Yves Oudeyer wrote down the article "Quand l'IA explore les pr  misses d'une vie artificielle" published in Pour La Science n  562 (August 2024).

PY Oudeyer wrote a note for the French educational institutions on "IA g  n  rative, soci  t   et   ducation: En quoi l'IA g  n  rative repr  sente-elle un enjeu dans la formation des citoyens ?", in the context of the *Conf  rence de Consensus on Nouveaux Savoirs et Nouvelles Comp  tences des Jeunes* of Cnesco, (Nov. 2024)

H  l  ne Sauz  on participated to the editing of 2 Popularization articles for the communication services of the centre Inria of University of UB : IA et apprentissage ; r  inventer l'  ducation).

H  l  ne Sauz  on participated to the "1 minute with..." video program by the popularization service : 1 Minute avec H. Sauz  on

### 11.3.3 Participation in Live events

Marion Pech & Ikram Chraibi Kaadoud (2024) « ChatGPT sera-t-il votre psychologue 2.0 ? ». Oral communication for the French Tech, November 21, Bordeaux. France.

Marion Pech, Florence Amardeilh, Gala Leibar, Trang Pham [DigitEco] (2024) Placéco program produced by Digital Aquitaine on the TV7 plateau in Sud-Ouest. La place des femmes dans la tech : le compte n'y est pas, comment faire ? Hosted by Gaëlle Richard. November 6, Bordeaux. France

Hélène Sauzéron took part in a round-table discussion; TABLE RONDE - Communautés, espaces sociaux, impact de l'IA sur la restructuration des sociétés Forum NAIA -Robocup, 2024, Bordeaux.

Hélène Sauzéron took part in a round-table discussion about the Education and generative AI, Colloque Infiné "Éducation et IA 11 octobre 2024, Futuroscope, Poitiers.

Hélène Sauzéron took part in the science popularization event for secondary schools Têtes chercheuses 2024.

Marion Pech, Florence Amardeilh, Gala Leibar, Trang Pham [DigitEco] (2024) Placéco program produced by Digital Aquitaine on the TV7 plateau in Sud-Ouest. La place des femmes dans la tech : le compte n'y est pas, comment faire ? Hosted by Gaëlle Richard. November 6, Bordeaux. France

Marion Pech (2024) « ChatGPT sera-t-il votre psychologue 2.0 ? ». Oral communication for the French Tech, November 21, Bordeaux. France.

PY Oudeyer participated to several live events:

- (Sept. 2024) **Presentation and interview** about the popular science book "C'est (pas) moi c'est l'IA" at an event organized by Nathan Jeunesse, Paris.
- (Oct. 2024) **Conceptualizing education and training alongside generative AI**, BIG Conference, organized by BPI.
- (Dec 2024) **Presentation and debate** about the topic "Comment éduquer à un usage responsable des IA ?", organized by CCIC-Unesco, Apprentis d'Auteuil.

#### 11.3.4 Others science outreach relevant activities

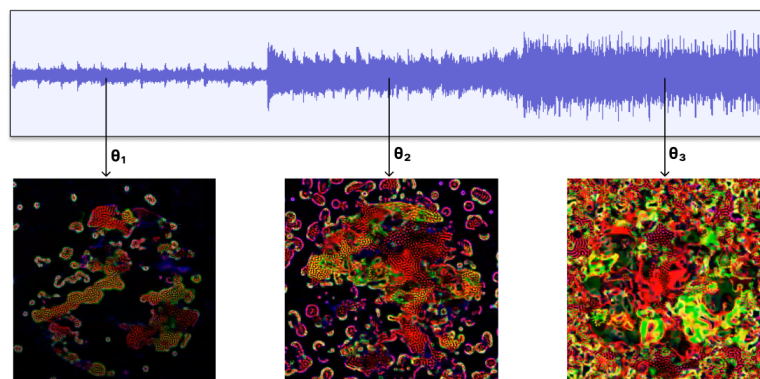
Clément Moulin-Frier co-organized the 3rd edition of Hack1robo with Xavier Hinaut (Mnemosyne), David Trocellier (Potioc) and Anne-Lise Pernel (SAER). The event took place at Le Node in Bordeaux in November 2024 (15-17) and received financial support from Inria, Inria Startup Studio, R3NumEd, R4 and GDR Robotique. <https://sites.google.com/view/hack1robo/accueil>. There was more than 40 participants, from teenagers to seniors (but mostly University students), working on 7 collaborative projects. Two of these projects were led by students of the Flowers team. The first project (led by Guillaume Pourcel, Julien Pourcel, Loris Gaven and Corentin Léger from Flowers) explored whether making Large Language Models more socially persuasive (rather than purely truth-seeking) could improve their reasoning abilities, by having them engage in group debates and testing if enhanced persuasiveness leads to better argument generation and evaluation. This project received the 1st prize of Hack1robo. The second project (led by Sina Khajehabdollahi, Jeremy Perez and Marko Cjvetko) merged music with visually representable dynamical systems to create music-driven animations. The core idea was to extract musical features—such as tempo, frequency, and other parameters—and use them to manipulate a dynamical system in real-time. For this purpose, the team selected Lenia, a class of continuous cellular automata renowned for generating diverse, life-like patterns. By mapping musical features to Lenia's parameters, they created a visually engaging, ever-evolving display of music-inspired animation. The implementation was built using TouchDesigner, a visual programming language for creating interactive multimedia content. This framework allows users to easily modify the interaction between music and Lenia, enabling efficient exploration of possible mappings. An example of how the audio signal and Lenia interact is shown in 55.

Hélène Sauzéron gave a talk on the sciences of autonomy and disability to the National Assembly's "Longevity and adaptation of society to ageing" study group, 2024;

Cécile Mazon participated to debates about teaching Cognitive Sciences during the CogGames (November 18th & 19th, Bordeaux).

Hélène Sauzéron gave an interview for the Sud-Ouest » news papers Sud-Ouest "Assistants vocaux : faut-il laisser les enfants utiliser seuls une enceinte connectée ? ;

Hélène Sauzéron provides the Scientific Council with support for innovation in companies with 2 actions in 2024 : for TelecomDesign and produit Vibby (contract being set up with STIP Bdx) and REVLIM via the European DIHNAMIC Nvelle Aquitaine - STIP Bdx project.



**Figure 55:** Example of a Lenia run modulated by an audio signal. At each timestep, Lenia parameters  $\theta$  are modulated by features extracted from the audio signal, altering the system's dynamics in response to the music.

PY. Oudeyer was interviewed, or the work of the team was discussed, in various newspapers, magazines and radios/podcasts:

- Scientific American, Dec. 2024, [Hitting the curiosity sweep spot speeds up learning](#).
- Famille et Education, Le magazine de l'APEL, Déc. 2024, [IA, un vrai ou faux ami?](#)
- Lemonde.fr, Oct. 2024, [MIA Seconde : à l'heure de ChatGPT, le tuteur sous intelligence artificielle de l'éducation nationale est-il déjà obsolète ?](#)
- Les Échos, Nov. 2024, [La France face au défi de la formation à l'IA](#).
- RFI, Sept. 2024, [Comment se familiariser avec l'IA ? Emission 8 milliards de voisins](#)
- France Culture, Sept. 2024, ["C'est pas moi, c'est l'IA", un manuel indispensable pour comprendre l'intelligence artificielle](#)
- Le Café Pédagogique, 2024, ["Il est urgent de former les profs à l'IA"](#)
- Le Monde Numérique" (Jérôme Colombain): ["Apprivoiser l'IA à l'école"](#)
- Blog Binaire (Lemonde.fr) : [Mon enfant apprivoise l'IA !](#)
- Sqool TV, Sept. 2024, [L'école du futur: comment apprivoiser l'IA?](#)
- La Nouvelle République, 2024, [L'IA, une aide aux devoirs?](#)
- Geek Junior: [C'est par moi c'est l'IA: Voici un guide malin et bienveillant pour découvrir comment fonctionne l'IA et à quoi elle sert](#)

## 12 Scientific production

### 12.1 Major publications

- [1] R. Abdelghani, P.-Y. Oudeyer, E. Law, C. de Vulpillières and H. Sauzéon. 'Conversational agents for fostering curiosity-driven learning in children'. In: *International Journal of Human-Computer Studies* 167 (Nov. 2022), p. 102887. DOI: [10.1016/j.ijhcs.2022.102887](https://doi.org/10.1016/j.ijhcs.2022.102887). URL: <https://hal.archives-ouvertes.fr/hal-03779096>.
- [2] M. Adolphe, M. Sawayama, D. Maurel, A. Delmas, P.-Y. Oudeyer and H. Sauzeon. 'An Open-Source Cognitive Test Battery to Assess Human Attention and Memory'. In: *Frontiers in Psychology* 13 (10th June 2022). DOI: [10.3389/fpsyg.2022.880375](https://doi.org/10.3389/fpsyg.2022.880375). URL: <https://hal.archives-ouvertes.fr/hal-03723887> (cit. on p. 66).

- [3] A. Akakzia, C. Colas, P.-Y. Oudeyer, M. Chetouani and O. Sigaud. ‘Grounding Language to Autonomously-Acquired Skills via Goal Generation’. In: *ICLR 2021 - Ninth International Conference on Learning Representation*. Vienna / Virtual, Austria, 4th May 2021. URL: <https://hal.inria.fr/hal-03121146>.
- [4] M. Alaimi, E. Law, K. D. Pantasdo, P.-Y. Oudeyer and H. Sauzéon. ‘Pedagogical Agents for Fostering Question-Asking Skills in Children’. In: *CHI '20 - CHI Conference on Human Factors in Computing Systems*. Honolulu / Virtual, United States, Apr. 2020. DOI: [10.1145/3313831.3376776](https://doi.org/10.1145/3313831.3376776). URL: <https://hal.inria.fr/hal-03109511>.
- [5] A. Baranes and P.-Y. Oudeyer. ‘Active Learning of Inverse Models with Intrinsically Motivated Goal Exploration in Robots’. In: *Robotics and Autonomous Systems* 61.1 (Jan. 2013), pp. 69–73. DOI: [10.1016/j.robot.2012.05.008](https://doi.org/10.1016/j.robot.2012.05.008). URL: <https://hal.inria.fr/hal-00788440>.
- [6] H. Caselles-Dupré, M. Garcia-Ortiz and D. Filliat. ‘S-TRIGGER: Continual State Representation Learning via Self-Triggered Generative Replay’. In: *IJCNN 2021 - International Joint Conference on Neural Networks*. Shenzhen / Virtual, China: IEEE, 18th July 2021, pp. 1–7. DOI: [10.1109/IJCNN52387.2021.9533683](https://doi.org/10.1109/IJCNN52387.2021.9533683). URL: <https://hal.archives-ouvertes.fr/hal-03377783>.
- [7] H. Caselles-Dupré, M. Garcia-Ortiz and D. Filliat. ‘Symmetry-Based Disentangled Representation Learning requires Interaction with Environments’. In: *NeurIPS 2019*. Vancouver, Canada, Dec. 2019. URL: <https://hal.archives-ouvertes.fr/hal-02379399>.
- [8] P.-A. Cinquin, P. Guitton and H. Sauzéon. ‘Towards Truly Accessible MOOCs for Persons with Cognitive Impairments: a Field Study’. In: *Human-Computer Interaction* (2021). URL: <https://hal.archives-ouvertes.fr/hal-03436244>.
- [9] C. Colas, P. Fournier, O. Sigaud, M. Chetouani and P.-Y. Oudeyer. ‘CURIOUS: Intrinsically Motivated Modular Multi-Goal Reinforcement Learning’. In: *International Conference on Machine Learning*. Long Beach, France, June 2019. URL: <https://hal.archives-ouvertes.fr/hal-01934921>.
- [10] C. Colas, B. P. Hejblum, S. Rouillon, R. Thiébaud, P.-Y. Oudeyer, C. Moulin-Frier and M. Prague. ‘EpidemiOptim: a Toolbox for the Optimization of Control Policies in Epidemiological Models’. In: *Journal of Artificial Intelligence Research* (23rd July 2021). DOI: [10.1613/jair.1.12588](https://doi.org/10.1613/jair.1.12588). URL: <https://hal.inria.fr/hal-03099898>.
- [11] C. Colas, T. Karch, N. Lair, J.-M. Dussoux, C. Moulin-Frier, P. F. Dominey and P.-Y. Oudeyer. ‘Language as a Cognitive Tool to Imagine Goals in Curiosity-Driven Exploration’. In: *NeurIPS 2020 - 34th Conference on Neural Information Processing Systems*. Contains main article and supplementaries. Vancouver / Virtual, Canada, Dec. 2020. URL: <https://hal.archives-ouvertes.fr/hal-03083158> (cit. on p. 11).
- [12] C. Colas, O. Sigaud and P.-Y. Oudeyer. ‘GEP-PG: Decoupling Exploration and Exploitation in Deep Reinforcement Learning Algorithms’. In: *International Conference on Machine Learning (ICML)*. Stockholm, Sweden, July 2018. URL: <https://hal.inria.fr/hal-01890151>.
- [13] C. Craye, T. Lesort, D. Filliat and J.-F. Goudou. ‘Exploring to learn visual saliency: The RL-IAC approach’. In: *Robotics and Autonomous Systems* 112 (Feb. 2019), pp. 244–259. URL: <https://hal.archives-ouvertes.fr/hal-01959882>.
- [14] N. Duminy, S. M. Nguyen, J. Zhu, D. Duhaut and J. Kerdreux. ‘Intrinsically Motivated Open-Ended Multi-Task Learning Using Transfer Learning to Discover Task Hierarchy’. In: *Applied Sciences* 11.3 (Feb. 2021), p. 975. DOI: [10.3390/app11030975](https://doi.org/10.3390/app11030975). URL: <https://hal.inria.fr/hal-03118190>.
- [15] M. Eppe and P.-Y. Oudeyer. ‘Intelligent Behavior Depends on the Ecological Niche’. In: *KI - Künstliche Intelligenz* (3rd Jan. 2021). DOI: [10.1007/s13218-020-00696-1](https://doi.org/10.1007/s13218-020-00696-1). URL: <https://hal.inria.fr/hal-03120583>.
- [16] M. Etcheverry, C. Moulin-Frier and P.-Y. Oudeyer. ‘Hierarchically Organized Latent Modules for Exploratory Search in Morphogenetic Systems’. In: *NeurIPS 2020 - 34th Conference on Neural Information Processing Systems*. Vancouver / Virtual, Canada, Dec. 2020. URL: <https://hal.inria.fr/hal-03099906> (cit. on pp. 74, 76).

- [17] S. Forestier, R. Portelas, Y. Mollard and P.-Y. Oudeyer. ‘Intrinsically Motivated Goal Exploration Processes with Automatic Curriculum Learning’. In: *Journal of Machine Learning Research* (1st Apr. 2022). URL: <https://hal.science/hal-01651233> (cit. on p. 76).
- [18] J. Gottlieb and P.-Y. Oudeyer. ‘Towards a neuroscience of active sampling and curiosity’. In: *Nature Reviews Neuroscience* 19.12 (Dec. 2018), pp. 758–770. URL: <https://hal.inria.fr/hal-01965608> (cit. on p. 5).
- [19] T. Karch, L. Teodorescu, K. Hofmann, C. Moulin-Frier and P.-Y. Oudeyer. ‘Grounding Spatio-Temporal Language with Transformers’. In: *NeurIPS 2021 - 35th Conference on Neural Information Processing Systems*. Virtuel, France, 7th Dec. 2021. URL: <https://hal.archives-ouvertes.fr/hal-03439096>.
- [20] A. Laversanne-Finot, A. Péré and P.-Y. Oudeyer. ‘Curiosity Driven Exploration of Learned Disentangled Goal Spaces’. In: *CoRL 2018 - Conference on Robot Learning*. Zürich, Switzerland, Oct. 2018. URL: <https://hal.inria.fr/hal-01891598>.
- [21] T. Lesort, N. Díaz-Rodríguez, J.-F. Goudou and D. Filliat. ‘State Representation Learning for Control: An Overview’. In: *Neural Networks* 108 (Dec. 2018), pp. 379–392. DOI: [10.1016/j.neunet.2018.07.006](https://doi.org/10.1016/j.neunet.2018.07.006). URL: <https://hal.archives-ouvertes.fr/hal-01858558>.
- [22] C. Mazon, B. Clément, D. Roy, P.-Y. Oudeyer and H. Sauzéon. ‘Pilot study of an intervention based on an intelligent tutoring system (ITS) for instructing mathematical skills of students with ASD and/or ID’. In: *Education and Information Technologies* (2022). DOI: [10.1007/s10639-022-11129-x](https://doi.org/10.1007/s10639-022-11129-x). URL: <https://hal.inria.fr/hal-03676589>.
- [23] M. E. Meade, J. G. Meade, H. Sauzéon and M. A. Fernandes. ‘Active Navigation in Virtual Environments Benefits Spatial Memory in Older Adults’. In: *Brain Sciences* 9 (2019). DOI: [10.3390/brainsci9030047](https://doi.org/10.3390/brainsci9030047). URL: <https://hal.inria.fr/hal-02049031>.
- [24] C. Moulin-Frier, J. Brochard, F. Stulp and P.-Y. Oudeyer. ‘Emergent Jaw Predominance in Vocal Development through Stochastic Optimization’. In: *IEEE Transactions on Cognitive and Developmental Systems* 99 (2017), pp. 1–12. DOI: [10.1109/TCDS.2017.2704912](https://doi.org/10.1109/TCDS.2017.2704912). URL: <https://hal.inria.fr/hal-01578075>.
- [25] E. Nisioti, K. Jodogne-del Litto and C. Moulin-Frier. ‘Grounding an Ecological Theory of Artificial Intelligence in Human Evolution’. In: *NeurIPS 2021 - Conference on Neural Information Processing Systems / Workshop: Ecological Theory of Reinforcement Learning*. virtual event, France, 6th Dec. 2021. URL: <https://hal.archives-ouvertes.fr/hal-03446961>.
- [26] A. Péré, S. Forestier, O. Sigaud and P.-Y. Oudeyer. ‘Unsupervised Learning of Goal Spaces for Intrinsically Motivated Goal Exploration’. In: *ICLR2018 - 6th International Conference on Learning Representations*. Vancouver, Canada, Apr. 2018. URL: <https://hal.archives-ouvertes.fr/hal-01891758>.
- [27] R. Portelas, C. Colas, K. Hofmann and P.-Y. Oudeyer. ‘Teacher algorithms for curriculum learning of Deep RL in continuously parameterized environments’. In: *CoRL 2019 - Conference on Robot Learning*. <https://arxiv.org/abs/1910.07224>. Osaka, Japan, Oct. 2019. URL: <https://hal.archives-ouvertes.fr/hal-02370165>.
- [28] R. Portelas, C. Colas, L. Weng, K. Hofmann and P.-Y. Oudeyer. ‘Automatic Curriculum Learning For Deep RL: A Short Survey’. In: *IJCAI 2020 - International Joint Conference on Artificial Intelligence*. Kyoto / Virtuelle, Japan, Jan. 2021. URL: <https://hal.archives-ouvertes.fr/hal-02923635>.
- [29] C. Reinke, M. Etcheverry and P.-Y. Oudeyer. ‘Intrinsically Motivated Discovery of Diverse Patterns in Self-Organizing Systems’. In: *International Conference on Learning Representations (ICLR)*. Source code and videos at <https://automated-discovery.github.io/>. Addis Ababa, Ethiopia, Apr. 2020. URL: <https://hal.inria.fr/hal-02370003> (cit. on pp. 74, 76).

- [30] C. Romac, R. Portelas, K. Hofmann and P.-Y. Oudeyer. ‘TeachMyAgent: a Benchmark for Automatic Curriculum Learning in Deep RL’. In: *Proceedings of the 38th International Conference on Machine Learning, PMLR 139, 2021*. ICML 2021 - Thirty-eighth International Conference on Machine Learning. Vol. 139. Proceedings of the 38th International Conference on Machine Learning. Vienna / Virtual, Austria, 18th July 2021, pp. 9052–9063. URL: <https://hal.archives-ouvertes.fr/hal-03173198>.
- [31] A. Ten, P. Kaushik, P.-Y. Oudeyer and J. Gottlieb. ‘Humans monitor learning progress in curiosity-driven exploration’. In: *Nature Communications* 12.1 (Dec. 2021). DOI: [10.1038/s41467-021-26196-w](https://doi.org/10.1038/s41467-021-26196-w). URL: <https://hal.inria.fr/hal-03378913> (cit. on p. 5).
- [32] G. Valle Perez, J. Beskow, G. E. Henter, A. Holzapfel, P.-Y. Oudeyer and S. Alexanderson. ‘Transflower: probabilistic autoregressive dance generation with multimodal attention’. In: SIGGRAPH Asia 2021 - 14th ACM SIGGRAPH Conference and Exhibition on Computer Graphics and Interactive Techniques. Tokyo, Japan, 14th Dec. 2021. DOI: [10.1145/3478513.3480570](https://doi.org/10.1145/3478513.3480570). URL: <https://hal.archives-ouvertes.fr/hal-03449915>.

## 12.2 Publications of the year

### International journals

- [33] M. Etcheverry, C. Moulin-Frier, P.-Y. Oudeyer and M. Levin. ‘AI-driven Automated Discovery Tools Reveal Diverse Behavioral Competencies of Biological Networks’. In: *eLife* (21st Aug. 2024). DOI: [10.7554/eLife.92683.3](https://doi.org/10.7554/eLife.92683.3). URL: <https://hal.science/hal-04395732> (cit. on pp. 14, 74–76).
- [34] H. Fournier, M. Hasdenteufel, C. Garrousteigt, M. Perie, A. Gliksohn, B. Jouanne, S. Hadj-Rabia, B. Arveiler, F. Morice-Picard and B. Quintard. ‘The experience of albinism in France: a qualitative study on dyads of parents and their adult child with albinism’. In: *BMC Medicine* 22.1 (29th Jan. 2024), p. 40. DOI: [10.1186/s12916-024-03251-z](https://doi.org/10.1186/s12916-024-03251-z). URL: <https://hal.science/hal-04501801>.
- [35] G. Kovač, R. Portelas, P. F. Dominey and P.-Y. Oudeyer. ‘The SocialAI school: a framework leveraging developmental psychology toward artificial socio-cultural agents’. In: *Frontiers in Neurorobotics* 18 (9th Oct. 2024). DOI: [10.3389/fnbot.2024.1396359](https://doi.org/10.3389/fnbot.2024.1396359). URL: <https://inria.hal.science/hal-04861861> (cit. on p. 13).
- [36] G. Kovač, R. Portelas, M. Sawayama, P. F. Dominey and P.-Y. Oudeyer. ‘Stick to your role! Stability of personal values expressed in large language models’. In: *PLoS ONE* 19.8 (26th Aug. 2024), e0309114. DOI: [10.1371/journal.pone.0309114](https://doi.org/10.1371/journal.pone.0309114). URL: <https://inria.hal.science/hal-04861867> (cit. on p. 13).
- [37] M. Poupard, F. Larrue, H. Sauzéon and A. Tricot. ‘A systematic review of immersive technologies for education: effects of cognitive load and curiosity state on learning performance’. In: *British Journal of Educational Technology* 56.1 (2025), pp. 5–41. DOI: [10.1111/bjet.13503](https://doi.org/10.1111/bjet.13503). URL: <https://hal.science/hal-03906797> (cit. on p. 71).
- [38] J. Pourcel, C. Colas, G. Molinaro, P.-Y. Oudeyer and L. Teodorescu. ‘ACES: Generating diverse programming puzzles with autotelic language models and semantic descriptors’. In: *Neurips* (2024). URL: <https://inria.hal.science/hal-04372580> (cit. on p. 82).
- [39] M. Sánchez-Fibla, C. Moulin-Frier and R. Solé. ‘Cooperative control of environmental extremes by artificial intelligent agents’. In: *Journal of the Royal Society Interface* 21.220 (2024). DOI: [10.48550/arXiv.2212.02395](https://doi.org/10.48550/arXiv.2212.02395). URL: <https://inria.hal.science/hal-04356920> (cit. on pp. 48, 84).
- [40] Y. Sivashankar, P. L. He, H. Sauzeon and M. A. Fernandes. ‘Motoric engagement, but not decision-making, during encoding influences route memory’. In: *Spatial Cognition and Computation* (30th May 2024), pp. 1–26. DOI: [10.1080/13875868.2024.2359929](https://doi.org/10.1080/13875868.2024.2359929). URL: <https://hal.science/hal-04673600>.
- [41] Y. Sivashankar, M. Fernandes, P.-Y. Oudeyer and H. Sauzéon. ‘The beneficial role of curiosity on route memory in children’. In: *Frontiers in Cognition* 3 (14th Mar. 2024). DOI: [10.3389/fcogn.2024.1346280](https://doi.org/10.3389/fcogn.2024.1346280). URL: <https://inria.hal.science/hal-04505921>.

**International peer-reviewed conferences**

- [42] L. Annabi, Z. Ma and S. M. Nguyen. ‘Unsupervised Motion Retargeting for Human-Robot Imitation’. In: HRI 2024 - Companion of the 2024 ACM/IEEE International Conference on Human-Robot Interaction. Boulder (CO), United States: ACM, 11th Mar. 2024. DOI: [10.1145/3568294.3580153](https://doi.org/10.1145/3568294.3580153). URL: <https://hal.science/hal-04401885>.
- [43] D. Bouchabou and S. M. Nguyen. ‘Generative Pretrained Embedding and Hierarchical Irregular Time Series Representation for Daily Living Activity Recognition’. In: *Frontiers in Artificial Intelligence and Applications*. ECAI 2024. Frontiers in Artificial Intelligence and Applications. Santiago de Compostela, Spain: IOS Press, 16th Oct. 2024, pp. 4764–4771. DOI: [10.3233/faia241075](https://doi.org/10.3233/faia241075). URL: <https://inria.hal.science/hal-04858744>.
- [44] D. Bouchabou, S. M. Nguyen, C. Lohr, B. Leduc and I. Kanellos. ‘Representation of Irregularly Sampled Time Series with Generative Language Models for Classification and Transfer Learning: a Case Study in Activities of Daily Living Recognition’. In: ML4ITS 2023 - Machine Learning for Irregular Time Series Workshop in ECML PKDD. Turin (IT), Italy, 17th Feb. 2024. URL: <https://hal.science/hal-04437409>.
- [45] D. Brellmann, E. Berthier, D. Filliat and G. Frehse. ‘On double descent in reinforcement learning with LSTD and random features’. In: *Proceedings of the Twelfth International Conference on Learning Representations (ICLR)*. ICLR 2024 - The Twelfth International Conference on Learning Representations. Vienna, Austria, 7th May 2024. URL: <https://hal.ip-paris.fr/hal-04628067>.
- [46] G. Franchi, M. Hariat, X. Yu, N. Belkhir, A. Manzanera and D. Filliat. ‘InfraParis: A multi-modal and multi-task autonomous driving dataset’. In: WACV 2024 - IEEE/CVF Winter Conference on Applications of Computer Vision. Waikoloa, United States, 2024. URL: <https://hal.science/hal-04321062>.
- [47] G. Kovač, R. Portelas, M. Sawayama, P. Ford Dominey and P.-Y. Oudeyer. ‘Stick to your Role! Stability of Personal Values Expressed in Large Language Models’. In: *Proceedings of the Annual Meeting of the Cognitive Science Society Volume 46*. CogSci 2024. Rotterdam, Netherlands, 24th July 2024. URL: <https://inria.hal.science/hal-04861881> (cit. on p. 37).
- [48] C. Léger, G. Hamon, E. Nisioti, X. Hinaut and C. Moulin-Frier. ‘Evolving Reservoirs for Meta Reinforcement Learning’. In: *Applications of Evolutionary Computation*. International Conference on the Applications of Evolutionary Computation (Part of EvoStar). Vol. 14635. Lecture Notes in Computer Science. Aberystwyth, United Kingdom: Springer Nature Switzerland, 2024, pp. 36–60. DOI: [10.1007/978-3-031-56855-8\\_3](https://doi.org/10.1007/978-3-031-56855-8_3). URL: <https://inria.hal.science/hal-04354303> (cit. on p. 51).
- [49] A. Marusic, S. M. Nguyen and A. Tapus. ‘PhysioFormer: A Spatio-Temporal Transformer for Physical Rehabilitation Assessment’. In: ICSR 2024. HRI ’23: Companion of the 2023 ACM/IEEE International Conference on Human-Robot Interaction. Odense, Denmark: Lecture Notes in Artificial Intelligence series (LNAI), 23rd Oct. 2024. DOI: [10.1145/3568294.3580153](https://doi.org/10.1145/3568294.3580153). URL: <https://hal.science/hal-04857956>.
- [50] G. Molinaro, C. Colas, P.-Y. Oudeyer and A. Collins. ‘Latent Learning Progress Drives Autonomous Goal Selection in Human Reinforcement Learning’. In: NeurIPS 2024 - The Thirty-Eighth Annual Conference on Neural Information Processing Systems. Vancouver, Canada, 8th June 2024. DOI: [10.31234/osf.io/5dtx3](https://doi.org/10.31234/osf.io/5dtx3). URL: <https://inria.hal.science/hal-04810193> (cit. on pp. 52, 82).
- [51] S. M. Nguyen, M. Devanne, O. Remy-Neris, M. Lempereur and A. Thepaut. ‘A Medical Low-Back Pain Physical Rehabilitation Dataset for Human Body Movement Analysis’. In: IJCNN 2024 - International Joint Conference on Neural Networks. Yokohama, Japan, 2024. URL: <https://inria.hal.science/hal-04629541>.
- [52] E. Nisioti, S. Risi, I. Momennejad, P.-Y. Oudeyer and C. Moulin-Frier. ‘Collective Innovation in Groups of Large Language Models’. In: ALIFE 2024 - The Conference on Artificial Life. Copenhagen, Denmark: MIT Press, 2024. DOI: [10.1162/isal\\_a\\_00730](https://doi.org/10.1162/isal_a_00730). URL: <https://inria.hal.science/hal-04848119> (cit. on pp. 13, 36).

- [53] G. Quere, F. Stulp, D. Filliat and J. Silverio. ‘A probabilistic approach for learning and adapting shared control skills with the human in the loop’. In: *Proceedings of the 2024 IEEE International Conference on Robotics and Automation*. ICRA 2024 - IEEE International Conference on Robotics and Automation. YOKOHAMA, Japan, 13th May 2024. URL: <https://hal.science/hal-04496397>.
- [54] M. Taylor-Davies, G. Hamon, T. Boulet and C. Moulin-Frier. ‘Emergent kin selection of altruistic feeding via non-episodic neuroevolution’. In: *Applications of Evolutionary Computation*. International Conference on the Applications of Evolutionary Computation (Part of EvoStar). Trieste, Italy, 2025. URL: <https://hal.science/hal-04865103> (cit. on p. 45).
- [55] M. Zadem, S. Mover, S. Mover and S. M. Nguyen. ‘Reconciling Spatial and Temporal Abstractions for Goal Representation’. In: ICLR 2024 - The Twelfth International Conference on Learning Representations. Vienna, Austria, 7th May 2024. URL: <https://inria.hal.science/hal-04629549>.

### Conferences without proceedings

- [56] Q. Gallouédec, E. Beeching, C. Romac and E. Dellandréa. ‘Jack of All Trades, Master of Some, a Multi-Purpose Transformer Agent’. In: ARLET 2024 - Aligning Reinforcement Learning Experimentalists and Theorists (Workshop at the International Conference on Machine Learning). Vienne, Austria, 26th July 2024. URL: <https://hal.science/hal-04681845>.
- [57] S. M. Nguyen. ‘Online Learning a Symbolic Abstraction of Actions in Hierarchical RL with Formal Methods’. In: ActSynt@ECAI2024. Santiago de Compostela, Spain, 19th Oct. 2024. URL: <https://inria.hal.science/hal-04858759>.
- [58] M. Poupard, F. Larrue, H. Sauzéon and A. Tricot. ‘The Interplay of Cognitive Load and Intrinsic Motivation in Neuroanatomy Learning through Immersive Technologies’. In: CLT 2024 - The 16th International Cognitive Load Theory Conference. Sydney, Australia, Nov. 2024. URL: <https://hal.science/hal-04816208> (cit. on p. 73).
- [59] G. Pourcel, T. Carta, G. Kovač and P.-Y. Oudeyer. ‘Autotelic LLM-based exploration for goal-conditioned RL’. In: Intrinsically Motivated Open-ended Learning Workshop at NeurIPS 2024. Vancouver, Canada, 15th Dec. 2024. URL: <https://inria.hal.science/hal-04861896>.

### Edition (books, proceedings, special issue of a journal)

- [60] *Ustensiles et espaces culinaires de la Protohistoire au début du XXe siècle, pré-actes du colloque international Corpus, Dijon, 4-7 juin 2024*. Ustensiles et espaces culinaires de la Protohistoire au début du XXe siècle. CORPUS, 4th June 2024, 79 p. URL: <https://hal.science/hal-04608861>.

### Doctoral dissertations and habilitation theses

- [61] M. Adolphe. ‘Development and evaluation of AI-based personalization algorithms for attention training’. Université de Bordeaux, 27th Sept. 2024. URL: <https://theses.hal.science/tel-04884647>.
- [62] S. M. Nguyen. ‘The intrinsic motivation of reinforcement and imitation learning for sequential tasks: Body movement analysis, activity of daily living analysis and intrinsically motivated robot for hierarchical reinforcement learning’. Sorbone Université, 27th Sept. 2024. URL: <https://hal.science/tel-04853270>.

### Reports & preprints

- [63] M. S. Aissi, C. Romac, T. Carta, S. Lamprier, P.-Y. Oudeyer, O. Sigaud, L. Soulier and N. Thome. *Reinforcement Learning for Aligning Large Language Models Agents with Interactive Environments: Quantifying and Mitigating Prompt Overfitting*. 17th Dec. 2024. URL: <https://hal.science/hal-04844077>.



- [64] B. Clément, H. Sauzéon, D. Roy and P.-Y. Oudeyer. *Improved Performances and Motivation in Intelligent Tutoring Systems: Combining Machine Learning and Learner Choice*. Inria Bordeaux Sud-Ouest, 1st Feb. 2024. URL: <https://inria.hal.science/hal-04433127>.
- [65] B. Clément, H. Sauzéon, D. Roy and P.-Y. Oudeyer. *Improved performances and motivation in intelligent tutoring systems: combining machine learning and learner choice*. 16th Jan. 2024. URL: <https://inria.hal.science/hal-04889350>.
- [66] G. Hamon, M. Etcheverry, B. W.-C. Chan, C. Moulin-Frier and P.-Y. Oudeyer. *Discovering Sensorimotor Agency in Cellular Automata using Diversity Search*. 14th Feb. 2024. URL: <https://hal.science/hal-04865117>.
- [67] J. Perez, G. Kovač, C. Léger, C. Colas, G. Molinaro, M. Derex, P.-Y. Oudeyer and C. Moulin-Frier. *When LLMs Play the Telephone Game: Cumulative Changes and Attractors in Iterated Cultural Transmissions*. 2024. URL: <https://hal.science/hal-04714994> (cit. on pp. 13, 40, 82).
- [68] J. Perez, C. Léger, M. Ovando-Tellez, C. Foulon, J. Dussauld, P.-Y. Oudeyer and C. Moulin-Frier. *Cultural evolution in populations of Large Language Models*. 2024. DOI: 10.48550/arXiv.2403.08882. URL: <https://hal.science/hal-04715009> (cit. on p. 39).
- [69] M. Poupard, F. Larrue, M. Bertrand, D. Liguoro, A. Tricot and H. Sauzéon. *Using virtual reality for enhancing neuroanatomy learning by optimizing cognitive load and intrinsic motivation*. Oct. 2024. URL: <https://hal.science/hal-04814452> (cit. on p. 71).
- [70] M. Srivastava, C. Colas, D. Sadigh and J. Andreas. *Policy Learning with a Language Bottleneck*. 2024. DOI: 10.48550/arXiv.2405.04118. URL: <https://inria.hal.science/hal-04810194>.
- [71] N. Yax, P.-Y. Oudeyer and S. Palminteri. *PhyloLM : Inferring the Phylogeny of Large Language Models and Predicting their Performances in Benchmarks*. 10th Jan. 2025. URL: <https://inria.hal.science/hal-04880495>.

#### Other scientific publications

- [72] L. Gaven, C. Romac, T. Carta, S. Lamprier, O. Sigaud and P.-Y. Oudeyer. ‘SAC-GLAM: Improving Online RL for LLM agents with Soft Actor-Critic and Hindsight Relabeling’. In: IMOL 2024 - Intrinsically Motivated Open-ended Learning (Workshop at Neurips). Vancouver, Canada, 16th Oct. 2024. URL: <https://hal.science/hal-04844089>.
- [73] G. Kovač, J. Perez, R. Portelas, M. Sawayama, P. F. Dominey and P.-Y. Oudeyer. *StickToYourRoleLeaderboard*. 2nd July 2024. URL: <https://inria.hal.science/hal-04870713> (cit. on p. 37).
- [74] M. Pech, M. Adolphe, P.-Y. Oudeyer and H. Sauzéon. ‘Benefits of Individualized Computerized Cognitive Training on older adults’. In: La Fédération pour la Recherche sur le Handicap et l’Autonomie (Fedrha). Bordeaux, France, 7th Nov. 2024. URL: <https://hal.science/hal-04848811>.
- [75] M. Poupard, F. Larrue, M. Bertrand, D. Liguoro, A. Tricot and H. Sauzéon. ‘Guided Drawing in Medical Education: Evaluating AR and MR Systems for Anatomy Learning’. In: STE 2024 - 4th International Conference on Science and Technology Education. Porto, Portugal, 3rd Oct. 2024. URL: <https://hal.science/hal-04816181> (cit. on p. 71).

#### Scientific popularization

- [76] C. Moulin-Frier, G. Hamon and P.-Y. Oudeyer. ‘Quand l’IA explore les prémices d’une vie artificielle’. In: *Pour la science* 562 (2024). URL: <https://inria.hal.science/hal-04848954>.
- [77] A. Nicole and M. Sbaraglia. *Digital Education Manual - Cycle 1 (3-4°) - Collection Décodage 2024*. Ed. by D. Roy. 1st Jan. 2024. URL: <https://inria.hal.science/hal-04473676>.
- [78] D. Roy, ed. *Education numérique - Manuel Cycle 2 (7-8°) - Collection Décodage 2024*. 8th Jan. 2024. URL: <https://inria.hal.science/hal-04410788>.

## Software

- [79] [SW] G. Hamon, *transformerXL\_PPO\_JAX* 24th July 2024. LIC: MIT License. HAL: ([hal-04659863](https://hal.inria.fr/hal-04659863)), URL: <https://inria.hal.science/hal-04659863>, VCS: [https://github.com/Reytuag/tranformerXL\\\_PPO\\\_JAX](https://github.com/Reytuag/tranformerXL\_PPO\_JAX).

## 12.3 Cited publications

- [80] M. Alaimi, E. Law, K. D. Pantasdo, P.-Y. Oudeyer and H. Sauz on. ‘Pedagogical Agents for Fostering Question-Asking Skills in Children’. In: CHI ’20 - CHI Conference on Human Factors in Computing Systems. Honolulu / Virtual, United States, 25th Apr. 2020. DOI: [10.1145/3313831.3376776](https://doi.org/10.1145/3313831.3376776). URL: <https://hal.inria.fr/hal-03109511> (cit. on p. 61).
- [81] A. Appriou, J. Ceha, S. Pramij, D. Dutartre, E. Law, P.-Y. Oudeyer and F. Lotte. ‘Towards measuring states of epistemic curiosity through electroencephalographic signals’. In: IEEE SMC 2020 - IEEE International conference on Systems, Man and Cybernetics. Toronto / Virtual, Canada, 11th Oct. 2020. URL: <https://hal.inria.fr/hal-02971795> (cit. on pp. 61, 62).
- [82] R. Abdelghani, E. Law, C. Desvaux, P.-Y. Oudeyer and H. Sauz on. ‘Interactive environments for training children’s curiosity through the practice of metacognitive skills : a pilot study’. In: *IDC 2023 - The 22nd annual ACM Interaction Design and Children Conference*. Chicago IL, United States: ACM, June 2023, pp. 495–501. DOI: [10.1145/3585088.3593880](https://doi.org/10.1145/3585088.3593880). URL: <https://hal.science/hal-04312937> (cit. on pp. 56, 58).
- [83] R. Abdelghani, P.-Y. Oudeyer, E. Law, C. de Vulpilli eres and H. Sauz on. ‘Conversational agents for fostering curiosity-driven learning in children’. In: *International Journal of Human-Computer Studies* 167 (Nov. 2022), p. 102887. DOI: [10.1016/j.ijhcs.2022.102887](https://doi.org/10.1016/j.ijhcs.2022.102887). URL: <https://hal.science/hal-03779096> (cit. on p. 54).
- [84] R. Abdelghani, H. Sauz on and P.-Y. Oudeyer. ‘Generative AI in the Classroom: Can Students Remain Active Learners?’ In: *NeurIPS 2023 - GAIED Workshop - Conference on Neural Information Processing Systems*. New orleans, USA, United States: arXiv, Dec. 2023. DOI: [10.48550/arXiv.2310.03192](https://doi.org/10.48550/arXiv.2310.03192). URL: <https://hal.science/hal-04312934> (cit. on p. 56).
- [85] R. Abdelghani, Y.-H. Wang, X. Yuan, T. Wang, P. Lucas, H. Sauz on and P.-Y. Oudeyer. ‘GPT-3-driven pedagogical agents for training children’s curious question-asking skills’. In: *International Journal of Artificial Intelligence in Education*. International Journal of Artificial Intelligence in Education (June 2023). DOI: [10.1007/s40593-023-00340-7](https://doi.org/10.1007/s40593-023-00340-7). URL: <https://hal.science/hal-03890447> (cit. on p. 54).
- [86] M. Adolphe, M. Pech, M. Sawayama, D. Maurel, A. Delmas, P.-Y. Oudeyer and H. Sauz on. ‘Exploring the Potential of Artificial Intelligence in Individualized Cognitive Training: a Systematic Review’. working paper or preprint. Dec. 2023. DOI: [10.31234/osf.io/2wg59](https://doi.org/10.31234/osf.io/2wg59). URL: <https://inria.hal.science/hal-04363997> (cit. on p. 65).
- [87] B. Argall, S. Chernova and M. Veloso. ‘A Survey of Robot Learning from Demonstration’. In: *Robotics and Autonomous Systems* 57.5 (2009), pp. 469–483 (cit. on p. 8).
- [88] M. Asada, S. Noda, S. Tawaratsumida and K. Hosoda. ‘Purposive Behavior Acquisition On A Real Robot By Vision-Based Reinforcement Learning’. In: *Machine Learning* 23 (1996), pp. 279–303 (cit. on p. 7).
- [89] A. Barto, S. Singh and N. Chentanez. ‘Intrinsically Motivated Learning of Hierarchical Collections of Skills’. In: *Proceedings of the 3rd International Conference on Development and Learning (ICDL 2004)*. Salk Institute, San Diego, 2004 (cit. on p. 8).
- [90] D. Berlyne. *Conflict, Arousal and Curiosity*. McGraw-Hill, 1960 (cit. on pp. 6, 8).
- [91] N. Bernstein. *The Coordination and Regulation of Movements*. Pergamon, 1967 (cit. on p. 5).
- [92] J. C. Brant and K. O. Stanley. ‘Minimal Criterion Coevolution: A New Approach to Open-Ended Search’. In: *Proceedings of the Genetic and Evolutionary Computation Conference*. GECCO ’17. 2017, pp. 67–74 (cit. on p. 45).

- [93] C. Breazeal. *Designing sociable robots*. The MIT Press, 2004 (cit. on p. 8).
- [94] R. Brooks, C. Breazeal, R. Irie, C. C. Kemp, B. Scassellati and M. Williamson. ‘Alternative essences of intelligence’. In: *Proceedings of 15th National Conference on Artificial Intelligence (AAAI-98)*. AAAI Press, 1998, pp. 961–968 (cit. on p. 7).
- [95] A. Cangelosi and M. Schlesinger. *Developmental robotics: From babies to robots*. MIT press, 2015 (cit. on p. 5).
- [96] T. Carta, C. Romac, T. Wolf, S. Lamprier, O. Sigaud and P.-Y. Oudeyer. ‘Grounding large language models in interactive environments with online reinforcement learning’. In: *arXiv preprint arXiv:2302.02662* (2023) (cit. on pp. 31, 32).
- [97] J. Ceha, E. Law, D. Kulić, v. Oudeyer Pierre-Y and D. Roy. ‘Identifying Functions and Behaviours of Social Robots for In-Class Learning Activities: Teachers’ Perspective’. In: *International Journal of Social Robotics* (Sept. 2021). DOI: [10.1007/s12369-021-00820-7](https://doi.org/10.1007/s12369-021-00820-7). URL: <https://hal.inria.fr/hal-03465477> (cit. on p. 61).
- [98] *Lenia and Expanded Universe*. Vol. ALIFE 2020: The 2020 Conference on Artificial Life. ALIFE 2021: The 2021 Conference on Artificial Life. July 2020, pp. 221–229. DOI: [10.1162/isal\\_a\\_00297](https://doi.org/10.1162/isal_a_00297). eprint: [https://direct.mit.edu/isal/proceedings-pdf/isal2020/32/221/1908612/isal\\_a\\_00297.pdf](https://direct.mit.edu/isal/proceedings-pdf/isal2020/32/221/1908612/isal_a_00297.pdf). URL: [https://doi.org/10.1162/isal%5C\\_a%5C\\_00297](https://doi.org/10.1162/isal%5C_a%5C_00297) (cit. on pp. 76, 78, 79).
- [99] B. W.-C. Chan. ‘Lenia-biology of artificial life’. In: *Complex Systems* 28.3 (2019), pp. 251–286 (cit. on pp. 76, 78, 79).
- [100] A. Clark. *Mindware: An Introduction to the Philosophy of Cognitive Science*. Oxford University Press, 2001 (cit. on p. 7).
- [101] B. Clément. ‘Adaptive Personalization of Pedagogical Sequences using Machine Learning’. Theses. Université de Bordeaux, Dec. 2018. URL: <https://hal.inria.fr/tel-01968241> (cit. on pp. 64, 85).
- [102] B. Clément, D. Roy, P.-Y. Oudeyer and M. Lopes. ‘Multi-Armed Bandits for Intelligent Tutoring Systems’. In: *Journal of Educational Data Mining (JEDM)* 7.2 (June 2015), pp. 20–48. URL: <https://hal.inria.fr/hal-00913669> (cit. on p. 64).
- [103] D. Cohn, Z. Ghahramani and M. Jordan. ‘Active learning with statistical models’. In: *Journal of artificial intelligence research* 4 (1996), pp. 129–145 (cit. on p. 8).
- [104] C. Colas, T. Karch, C. Moulin-Frier and P.-Y. Oudeyer. ‘Language and culture internalization for human-like autotelic AI’. In: *Nature Machine Intelligence* 4.12 (1st Dec. 2022), pp. 1068–1076. DOI: [10.1038/s42256-022-00591-4](https://doi.org/10.1038/s42256-022-00591-4). URL: <https://doi.org/10.1038/s42256-022-00591-4> (cit. on p. 26).
- [105] C. Colas, T. Karch, O. Sigaud and P.-Y. Oudeyer. ‘Autotelic agents with intrinsically motivated goal-conditioned reinforcement learning: a short survey’. In: *Journal of Artificial Intelligence Research* 74 (2022), pp. 1159–1199 (cit. on p. 26).
- [106] C. Colas, T. Karch, O. Sigaud and P.-Y. Oudeyer. ‘Intrinsically Motivated Goal-Conditioned Reinforcement Learning: a Short Survey’. working paper or preprint. Jan. 2021. URL: <https://hal.inria.fr/hal-03099891> (cit. on p. 76).
- [107] W. Croft and D. Cruse. *Cognitive Linguistics*. Cambridge Textbooks in Linguistics. Cambridge University Press, 2004 (cit. on p. 7).
- [108] M. Csikszentmihalyi. *Flow-the psychology of optimal experience*. Harper Perennial, 1991 (cit. on pp. 6, 8).
- [109] P. Dayan and W. Belleine. ‘Reward, motivation and reinforcement learning’. In: *Neuron* 36 (2002), pp. 285–298 (cit. on p. 8).
- [110] E. Deci and R. Ryan. *Intrinsic Motivation and Self-Determination in Human Behavior*. Plenum Press, 1985 (cit. on p. 8).

- [111] M. Derex and R. Boyd. ‘Partial connectivity increases cultural accumulation within groups’. en. In: *Proceedings of the National Academy of Sciences* 113.11 (Mar. 2016), pp. 2982–2987. DOI: [10.1073/pnas.1518798113](https://doi.org/10.1073/pnas.1518798113). URL: <http://www.pnas.org/lookup/doi/10.1073/pnas.1518798113> (visited on 29/01/2021) (cit. on p. 35).
- [112] Y. Duan, J. Schulman, X. Chen, P. L. Bartlett, I. Sutskever and P. Abbeel. ‘RL<sup>2</sup>: Fast Reinforcement Learning via Slow Reinforcement Learning’. In: *arXiv:1611.02779 [cs, stat]* (2016) (cit. on p. 51).
- [113] J. Elman. ‘Learning and development in neural networks: The importance of starting small’. In: *Cognition* 48 (1993), pp. 71–99 (cit. on p. 7).
- [114] M. Etcheverry. ‘Curiosity-driven AI for Science : Automated Discovery of Self-Organized Structures’. Theses. Université de Bordeaux, Nov. 2023. URL: <https://theses.hal.science/tel-04504878> (cit. on p. 14).
- [115] M. Etcheverry, M. Levin, C. Moulin-Frier and P.-Y. Oudeyer. ‘SBMLTOODEJAX: Efficient Simulation and Optimization of Biological Network Models in JAX’. In: *NeurIPS 2023 - AI4Science Workshop*. New Orleans (LA), France, Dec. 2023. URL: <https://hal.science/hal-04317246> (cit. on p. 74).
- [116] J. Gottlieb, P.-Y. Oudeyer, M. Lopes and A. Baranes. ‘Information-seeking, curiosity, and attention: computational and neural mechanisms’. In: *Trends in Cognitive Sciences* 17.11 (Nov. 2013), pp. 585–93. DOI: [10.1016/j.tics.2013.09.001](https://doi.org/10.1016/j.tics.2013.09.001). URL: <https://hal.inria.fr/hal-00913646> (cit. on p. 6).
- [117] J. Grizou, L. J. Points, A. Sharma and L. Cronin. ‘A curious formulation robot enables the discovery of a novel protocell behavior’. In: *Science Advances* 6.5 (2020), eaay4237. DOI: [10.1126/sciadv.aay4237](https://doi.org/10.1126/sciadv.aay4237). eprint: <https://www.science.org/doi/pdf/10.1126/sciadv.aay4237>. URL: <https://www.science.org/doi/abs/10.1126/sciadv.aay4237> (cit. on p. 10).
- [118] W. D. Hamilton. ‘The genetical evolution of social behaviour. I’. In: *Journal of Theoretical Biology* 7.1 (July 1964), pp. 1–16. DOI: [10.1016/0022-5193\(64\)90038-4](https://doi.org/10.1016/0022-5193(64)90038-4). URL: <https://www.sciencedirect.com/science/article/pii/0022519364900384> (cit. on p. 47).
- [119] W. Hamilton. ‘The genetical evolution of social behaviour. II’. In: *Journal of Theoretical Biology* 7.1 (1964), pp. 17–52. DOI: [https://doi.org/10.1016/0022-5193\(64\)90039-6](https://doi.org/10.1016/0022-5193(64)90039-6). URL: <https://www.sciencedirect.com/science/article/pii/0022519364900396> (cit. on p. 47).
- [120] G. Hamon, M. Etcheverry, B. W.-C. Chan, C. Moulin-Frier and P.-Y. Oudeyer. *Learning Sensorimotor Agency in Cellular Automata*. In this blogpost, we explore the concepts of embodiment, individuality, self-maintenance and sensorimotor agency within a cellular automaton (CA) environment. Whereas those concepts are central in theoretical biology and cognitive science, it remains unclear how such behaviors can emerge in a CA-like environment made only of low-level particles and physical rules. We present a novel set of tools (based on curriculum learning, diversity search and gradient descent over a differentiable CA) to automatically learn the rules leading to the emergence of such behaviors. Our method is able to discover robust self-organizing agents with strong coherence and generalization to out-of-distribution changes, reminiscent of the robustness of living systems to maintain specific functions despite environmental and body perturbations. Jan. 2022. URL: <https://inria.hal.science/hal-03519319> (cit. on p. 74).
- [121] S. Harnad. ‘The symbol grounding problem’. In: *Physica D* 40 (1990), pp. 335–346 (cit. on p. 7).
- [122] M. Hasenjager and H. Ritter. *Active learning in neural networks*. Heidelberg, Germany, Germany: Physica-Verlag GmbH, 2002, pp. 137–169 (cit. on p. 8).
- [123] J. Haugeland. *Artificial Intelligence: the very idea*. Cambridge, MA, USA: The MIT Press, 1985 (cit. on p. 7).
- [124] J.-C. Horvitz. ‘Mesolimbocortical and nigrostriatal dopamine responses to salient non-reward events’. In: *Neuroscience* 96.4 (2000), pp. 651–656 (cit. on p. 8).

- [125] X. Huang and J. Weng. ‘Novelty and reinforcement learning in the value system of developmental robots’. In: *Proceedings of the 2nd international workshop on Epigenetic Robotics : Modeling cognitive development in robotic systems*. Ed. by C. Prince, Y. Demiris, Y. Marom, H. Kozima and C. Balkenius. Lund University Cognitive Studies 94, 2002, pp. 47–55 (cit. on p. 8).
- [126] S. Ivaldi, N. Lyubova, D. G erardeaux-Viret, A. Droniou, S. Anzalone, M. Chetouani, D. Filliat and O. Sigaud. ‘Perception and human interaction for developmental learning of objects and affordances’. In: *Proc. of the 12th IEEE-RAS International Conference on Humanoid Robots - HUMANOIDS*. forthcoming. Japan, 2012. URL: <http://hal.inria.fr/hal-00755297> (cit. on p. 8).
- [127] M. Jaderberg, W. M. Czarnecki, I. Dunning, L. Marris, G. Lever, A. G. Castaneda, C. Beattie, N. C. Rabinowitz, A. S. Morcos, A. Ruderman et al. ‘Human-level performance in 3D multiplayer games with population-based reinforcement learning’. In: *Science* 364.6443 (2019). Publisher: American Association for the Advancement of Science, pp. 859–865 (cit. on p. 4).
- [128] M. Johnson. *Developmental Cognitive Neuroscience*. 2nd. Blackwell publishing, 2005 (cit. on p. 7).
- [129] M. H. Johnson. *Developmental cognitive neuroscience*. Wiley-Blackwell, 2011 (cit. on p. 5).
- [130] T. D. Johnston. ‘Selective Costs and Benefits in the Evolution of Learning’. en. In: *Advances in the Study of Behavior*. Ed. by J. S. Rosenblatt, R. A. Hinde, C. Beer and M.-C. Busnel. Vol. 12. Academic Press, Jan. 1982, pp. 65–106. DOI: [10.1016/S0065-3454\(08\)60046-7](https://doi.org/10.1016/S0065-3454(08)60046-7). URL: <http://www.sciencedirect.com/science/article/pii/S0065345408600467> (visited on 22/04/2020) (cit. on p. 51).
- [131] H. Kitano. ‘Biological robustness’. In: *Nature Reviews Genetics* 5.11 (2004), pp. 826–837 (cit. on p. 77).
- [132] W. B. Knox and P. Stone. ‘Combining manual feedback with subsequent MDP reward signals for reinforcement learning’. In: *Proceedings of the 9th International Conference on Autonomous Agents and Multiagent Systems (AAMAS’10)*. Toronto, Canada, 2010, pp. 5–12 (cit. on p. 8).
- [133] G. Kova , M. Sawayama, R. Portelas, C. Colas, P. F. Dominey and P.-Y. Oudeyer. ‘Large Language Models as Superpositions of Cultural Perspectives’. Preprint. Dec. 2023. URL: <https://inria.hal.science/hal-04365783> (cit. on p. 37).
- [134] K. N. Laland, T. Uller, M. W. Feldman, K. Sterelny, G. B. M ller, A. Moczek, E. Jablonka and J. Odling-Smee. ‘The extended evolutionary synthesis: its structure, assumptions and predictions’. In: *Proceedings of the royal society B: biological sciences* 282.1813 (2015), p. 20151019 (cit. on p. 44).
- [135] R. T. Lange and H. Sprekeler. *Learning not to learn: Nature versus nurture in silico*. 2020. arXiv: [2010.04466](https://arxiv.org/abs/2010.04466) [cs.LG] (cit. on p. 51).
- [136] J. Z. Leibo, V. Zambaldi, M. Lanctot, J. Marecki and T. Graepel. ‘Multi-Agent Reinforcement Learning in Sequential Social Dilemmas’. In: *Proceedings of the 16th Conference on Autonomous Agents and MultiAgent Systems. AAMAS ’17*. S o Paulo, Brazil, 2017, pp. 464–473 (cit. on p. 45).
- [137] M. Lopes, T. Cederborg and P.-Y. Oudeyer. ‘Simultaneous Acquisition of Task and Feedback Models’. In: *Development and Learning (ICDL), 2011 IEEE International Conference on*. Germany, 2011, pp. 1–7. DOI: [10.1109/DEVLRN.2011.6037359](https://doi.org/10.1109/DEVLRN.2011.6037359). URL: <http://hal.inria.fr/hal-00636166/en> (cit. on p. 8).
- [138] M. Lungarella, G. Metta, R. Pfeifer and G. Sandini. ‘Developmental Robotics: A Survey’. In: *Connection Science* 15.4 (2003), pp. 151–190 (cit. on pp. 5, 7).
- [139] N. Lyubova and D. Filliat. ‘Developmental Approach for Interactive Object Discovery’. In: *Neural Networks (IJCNN), The 2012 International Joint Conference on*. Australia, June 2012, pp. 1–7. DOI: [10.1109/IJCNN.2012.6252606](https://doi.org/10.1109/IJCNN.2012.6252606). URL: <https://hal.archives-ouvertes.fr/hal-00755298> (cit. on p. 8).
- [140] J. Marshall, D. Blank and L. Meeden. ‘An Emergent Framework for Self-Motivation in Developmental Robotics’. In: *Proceedings of the 3rd International Conference on Development and Learning (ICDL 2004)*. Salk Institute, San Diego, 2004 (cit. on p. 8).

- [141] M. Mason and M. Lopes. 'Robot Self-Initiative and Personalization by Learning through Repeated Interactions'. In: *6th ACM/IEEE International Conference on Human-Robot*. Switzerland, 2011. DOI: [10.1145/1957656.1957814](https://doi.org/10.1145/1957656.1957814). URL: <http://hal.inria.fr/hal-00636164/en> (cit. on p. 8).
- [142] C. Mazon, B. Clément, D. Roy, P.-Y. Oudeyer and H. Sauzéon. 'Pilot study of an intervention based on an intelligent tutoring system (ITS) for instructing mathematical skills of students with ASD and/or ID'. In: *Education and Information Technologies* (2022). DOI: [10.1007/s10639-022-11129-x](https://doi.org/10.1007/s10639-022-11129-x). URL: <https://inria.hal.science/hal-03676589> (cit. on p. 65).
- [143] C. Mazon, K. Etchegoyhen, I. Saint-Supery, A. Amestoy, M. Bouvard, C. Consel and H. Sauzéon. 'Fostering parents-professional collaboration for facilitating the school inclusion of students with ASD: Design of the "ToGather" web-based prototype'. In: *Educational Technology Research and Development* (Dec. 2021). DOI: [10.1007/s11423-021-10073-w](https://doi.org/10.1007/s11423-021-10073-w). URL: <https://hal.inria.fr/hal-03436355> (cit. on p. 70).
- [144] C. Mazon, C. Fage and H. Sauzéon. 'Effectiveness and usability of technology-based interventions for children and adolescents with ASD: A systematic review of reliability, consistency, generalization and durability related to the effects of intervention'. In: *Computers in Human Behavior* 93 (Apr. 2019). DOI: [10.1016/j.chb.2018.12.001](https://doi.org/10.1016/j.chb.2018.12.001). URL: <https://hal.inria.fr/hal-01950078> (cit. on p. 65).
- [145] C. Mazon and H. Sauzéon. 'Utilisation des technologies mobiles auprès des enfants avec TSA.' In: *Autisme et usages du numériques en éducation*. Ed. by P. Bourdon. 2022. URL: <https://inria.hal.science/hal-03120323> (cit. on p. 65).
- [146] E. Meyer, H. Sauzéon, I. Saint-Supery and C. Mazon. 'Systematic review of technologies to collaborate and co-educate students with special educational needs and supporting their schooling'. In: *IHIET 2023 - 10th International Conference on Human Interaction and Emerging Technologies*. Vol. 111. Nice, France: AHFE International, Aug. 2023, pp. 1–12. DOI: [10.54941/ahfe1004060](https://doi.org/10.54941/ahfe1004060). URL: <https://inria.hal.science/hal-04222434> (cit. on p. 70).
- [147] P. Miller. *Theories of developmental psychology*. 4th. New York: Worth, 2001 (cit. on pp. 5, 7, 8).
- [148] V. Mnih, K. Kavukcuoglu, D. Silver, A. A. Rusu, J. Veness, M. G. Bellemare, A. Graves, M. Riedmiller, A. K. Fidjeland, G. Ostrovski, S. Petersen, C. Beattie, A. Sadik, I. Antonoglou, H. King, D. Kumaran, D. Wierstra, S. Legg and D. Hassabis. 'Human-level control through deep reinforcement learning'. en. In: *Nature* 518.7540 (Feb. 2015), pp. 529–533. DOI: [10.1038/nature14236](https://doi.org/10.1038/nature14236). URL: <http://www.nature.com/articles/nature14236> (visited on 22/08/2018) (cit. on p. 4).
- [149] C. Moulin-Frier. 'The Ecology of Open-Ended Skill Acquisition'. Habilitation à diriger des recherches. Université de Bordeaux (UB), Dec. 2022. URL: <https://inria.hal.science/tel-03875448> (cit. on p. 43).
- [150] K. Murayama, L. FitzGibbon and M. Sakaki. 'Process Account of Curiosity and Interest: A Reward-Learning Perspective'. en. In: *Educational Psychology Review* 31.4 (Dec. 2019), pp. 875–895. DOI: [10.1007/s10648-019-09499-9](https://doi.org/10.1007/s10648-019-09499-9). URL: <http://link.springer.com/10.1007/s10648-019-09499-9> (visited on 04/10/2021) (cit. on p. 56).
- [151] A. Nair, P. Srinivasan, S. Blackwell, C. Alcicek, R. Fearon, A. De Maria, V. Panneershelvam, M. Suleyman, C. Beattie, S. Petersen, S. Legg, V. Mnih, K. Kavukcuoglu and D. Silver. *Massively Parallel Methods for Deep Reinforcement Learning*. en. Tech. rep. arXiv:1507.04296. arXiv:1507.04296 [cs]. arXiv, July 2015. URL: <http://arxiv.org/abs/1507.04296> (visited on 15/05/2022) (cit. on p. 35).
- [152] S. M. Nguyen, A. Baranes and P.-Y. Oudeyer. 'Bootstrapping Intrinsically Motivated Learning with Human Demonstrations'. In: *IEEE International Conference on Development and Learning*. Frankfurt, Germany, 2011. URL: <http://hal.inria.fr/hal-00645986/en> (cit. on p. 8).
- [153] S. M. Nguyen, A. Baranes and P.-Y. Oudeyer. 'Constraining the Size Growth of the Task Space with Socially Guided Intrinsic Motivation using Demonstrations.' In: *IJCAI Workshop on Agents Learning Interactively from Human Teachers (ALIHT)*. Barcelona, Spain, 2011. URL: <http://hal.inria.fr/hal-00645995/en> (cit. on p. 8).

- [154] E. Nisioti, M. Mahaut, P.-Y. Oudeyer, I. Momennejad and C. Moulin-Frier. ‘Social Network Structure Shapes Innovation: Experience-sharing in RL with SAPIENS’. working paper or preprint. July 2022. URL: <https://hal.science/hal-03715924> (cit. on p. 36).
- [155] P.-Y. Oudeyer. ‘Developmental Robotics’. In: *Encyclopedia of the Sciences of Learning*. Ed. by N. Seel. Springer Reference Series. Springer, 2011. URL: <http://hal.inria.fr/hal-00652123/en> (cit. on p. 5).
- [156] P.-Y. Oudeyer. ‘L’auto-organisation dans l’évolution de la parole’. In: *Parole et Musique: Aux origines du dialogue humain, Colloque annuel du Collège de France*. Ed. by S. Dehaene and C. Petit. Odile Jacob, 2009, pp. 83–112. URL: <http://hal.inria.fr/inria-00446908/en/> (cit. on p. 7).
- [157] P.-Y. Oudeyer. ‘Sur les interactions entre la robotique et les sciences de l’esprit et du comportement’. In: *Informatique et Sciences Cognitives : influences ou confluences ?* Ed. by C. Garbay and D. Kaiser. Presses Universitaires de France, 2009. URL: <http://hal.inria.fr/inria-00420309/en/> (cit. on p. 7).
- [158] P.-Y. Oudeyer, F. Kaplan and V. Hafner. ‘Intrinsic Motivation Systems for Autonomous Mental Development’. In: *IEEE Transactions on Evolutionary Computation* 11.2 (2007), pp. 265–286. DOI: [10.1109/TEVC.2006.89027](https://doi.org/10.1109/TEVC.2006.89027) (cit. on pp. 7, 8).
- [159] P.-Y. Oudeyer and F. Kaplan. ‘Intelligent adaptive curiosity: a source of self-development’. In: *Proceedings of the 4th International Workshop on Epigenetic Robotics*. Ed. by L. Berthouze, H. Kozima, C. Prince, G. Sandini, G. Stojanov, G. Metta and C. Balkenius. Vol. 117. Lund University Cognitive Studies, 2004, pp. 127–130 (cit. on p. 8).
- [160] P.-Y. Oudeyer and F. Kaplan. ‘What is intrinsic motivation? A typology of computational approaches’. In: *Frontiers in Neurorobotics* 1.1 (2007) (cit. on p. 8).
- [161] D. Pathak, P. Agrawal, A. A. Efros and T. Darrell. ‘Curiosity-driven exploration by self-supervised prediction’. In: *International conference on machine learning*. PMLR, 2017, pp. 2778–2787 (cit. on p. 31).
- [162] K. Pearson. ‘Neural adaptation in the generation of rhythmic behavior’. In: *Annual review of physiology* 62.1 (2000), pp. 723–753 (cit. on p. 51).
- [163] J. Perolat, J. Z. Leibo, V. Zambaldi, C. Beattie, K. Tuyls and T. Graepel. ‘A multi-agent reinforcement learning model of common-pool resource appropriation’. In: *Advances in neural information processing systems* 30 (2017) (cit. on pp. 44, 45).
- [164] E. Plantec, G. Hamon, M. Etcheverry, P.-Y. Oudeyer, C. Moulin-Frier and B. W.-C. Chan. ‘Flow-Lenia: Towards open-ended evolution in cellular automata through mass conservation and parameter localization’. In: *The 2023 Conference on Artificial Life*. Tokyo, Japan: MIT Press, July 2023. DOI: [10.1162/isal\\_a\\_00651](https://doi.org/10.1162/isal_a_00651). URL: <https://hal.science/hal-03904389> (cit. on p. 78).
- [165] S. Reed, K. Zolna, E. Parisotto, S. G. Colmenarejo, A. Novikov, G. Barth-maroon, M. Giménez, Y. Sulsky, J. Kay, J. T. Springenberg, T. Eccles, J. Bruce, A. Razavi, A. Edwards, N. Heess, Y. Chen, R. Hadsell, O. Vinyals, M. Bordbar and N. de Freitas. ‘A Generalist Agent’. In: *Transactions on Machine Learning Research* (2022). Featured Certification, Outstanding Certification (cit. on p. 33).
- [166] A. Revel and J. Nadel. ‘Imitation and Social Learning in Robots, Humans and Animals: Behavioural, Social and Communicative Dimensions’. In: ed. by K. Dautenhahn and C. Nehaniv. Cambridge University Press, 2004. Chap. How to build an imitator? (Cit. on p. 8).
- [167] I. Saint-Supery, C. Mazon, M. Eric and H. Sauzéon. ‘Conception d’une application de soutien à la coéducation pour l’inclusion scolaire des élèves TSA’. In: *Éthiques inclusives en éducation. Recherches, contextes et pratiques (p. 145-160)*. Parentalité & Handicap. Champs Social, 2023, p. 260. URL: <https://inria.hal.science/hal-04147610> (cit. on p. 70).
- [168] I. Saint-Supery, H. Sauzéon, C. Maillart, N. Neu, E. Meyer and C. Mazon. ‘Cross-cultural evaluation of a web application to support communication and collaboration among stakeholders of the school inclusion of children with ASD’. In: *AAATE 2023 - The 17th International Conference of the Association for the Advancement of Assistive Technology in Europe*. AAATE. Paris, France, Aug. 2023. URL: <https://inria.hal.science/hal-04147612> (cit. on p. 70).

- [169] I. Saint-Supéry, H. Sauzéon, E. Meyer and C. Mazon. ‘ToGather, an interactive website for the stakeholders of school inclusion of children with ASD: an iterative design including user testing’. working paper or preprint. 2022. URL: <https://inria.hal.science/hal-03898867> (cit. on p. 70).
- [170] T. Schatz and P-Y. Oudeyer. ‘Learning motor dependent Crutchfield’s information distance to anticipate changes in the topology of sensory body maps’. In: *IEEE International Conference on Learning and Development*. Chine Shangai, 2009. URL: <http://hal.inria.fr/inria-00420186/en/> (cit. on p. 8).
- [171] M. Schembri, M. Mirolli and G. Baldassarre. ‘Evolving internal reinforcers for an intrinsically motivated reinforcement-learning robot’. In: *IEEE 6th International Conference on Development and Learning, 2007. ICDL 2007*. (July 2007), pp. 282–287. URL: <http://dx.doi.org/10.1109/DEVLRN.2007.4354052> (cit. on p. 8).
- [172] J. Schmidhuber. ‘Curious Model-Building Control Systems’. In: *Proceedings of the International Joint Conference on Neural Networks, Singapore*. Vol. 2. IEEE press, 1991, pp. 1458–1463 (cit. on p. 8).
- [173] B. Schrauwen, D. Verstraeten and J. M. V. Campenhout. ‘An overview of reservoir computing: theory, applications and implementations’. In: *The European Symposium on Artificial Neural Networks*. 2007. URL: <https://api.semanticscholar.org/CorpusID:16935574> (cit. on p. 52).
- [174] W. Schultz, P. Dayan and P. Montague. ‘A neural substrate of prediction and reward’. In: *Science* 275 (1997), pp. 1593–1599 (cit. on p. 8).
- [175] I. Shumailov, Z. Shumaylov, Y. Zhao, N. Papernot, R. Anderson and Y. Gal. ‘AI models collapse when trained on recursively generated data’. In: *Nature* 631.8022 (July 2024), pp. 755–759. DOI: [10.1038/s41586-024-07566-y](https://doi.org/10.1038/s41586-024-07566-y). URL: <https://doi.org/10.1038/s41586-024-07566-y> (cit. on p. 40).
- [176] D. Silver, A. Huang, C. J. Maddison, A. Guez, L. Sifre, G. Van Den Driessche, J. Schrittwieser, I. Antonoglou, V. Panneershelvam, M. Lanctot et al. ‘Mastering the game of Go with deep neural networks and tree search’. In: *nature* 529.7587 (2016), pp. 484–489 (cit. on p. 4).
- [177] Y. Sivashankar, M. Fernandes, P-Y. Oudeyer and H. Sauzéon. ‘The Beneficial Role of Curiosity on Route memory in Children’. working paper or preprint. Jan. 2024. DOI: [10.31234/osf.io/4pdca](https://doi.org/10.31234/osf.io/4pdca). URL: <https://inria.hal.science/hal-04395366> (cit. on p. 62).
- [178] J. M. Smith. ‘Group selection and kin selection’. In: *Nature* 201 (1964), pp. 1145–1147. URL: <https://doi.org/10.1038/2011145a0> (cit. on p. 47).
- [179] K. O. Stanley, J. Clune, J. Lehman and R. Miikkulainen. ‘Designing Neural Networks through Neuroevolution’. In: *Nature Machine Intelligence* 1.1 (Jan. 2019), pp. 24–35. DOI: [10.1038/s42256-018-0006-z](https://doi.org/10.1038/s42256-018-0006-z) (cit. on p. 44).
- [180] L. Steels and R. Brooks, eds. *The Artificial Life Route to Artificial Intelligence: Building Embodied, Situated Agents*. Hillsdale, NJ, USA: L. Erlbaum Associates Inc., 1995 (cit. on p. 7).
- [181] D. W. Stephens. ‘Change, regularity, and value in the evolution of animal learning’. In: *Behavioral Ecology* 2.1 (1991), pp. 77–89. DOI: [10.1093/beheco/2.1.77](https://doi.org/10.1093/beheco/2.1.77) (cit. on p. 51).
- [182] E. Thelen and L. B. Smith. *A dynamic systems approach to the development of cognition and action*. Cambridge, MA: MIT Press, 1994 (cit. on p. 7).
- [183] A. L. Thomaz and C. Breazeal. ‘Teachable robots: Understanding human teaching behavior to build more effective robot learners’. In: *Artificial Intelligence Journal* 172 (2008), pp. 716–737 (cit. on p. 8).
- [184] A. Tierney. ‘Evolutionary implications of neural circuit structure and function’. In: *Behavioural processes* 35.1-3 (1995), pp. 173–182 (cit. on p. 51).
- [185] P. A. Tsividis, J. Loula, J. Burga, N. Foss, A. Campero, T. Pouncy, S. J. Gershman and J. B. Tenenbaum. *Human-Level Reinforcement Learning through Theory-Based Modeling, Exploration, and Planning*. 27th July 2021. arXiv: [2107.12544](https://arxiv.org/abs/2107.12544). URL: <http://arxiv.org/abs/2107.12544> (visited on 18/10/2024) (cit. on p. 31).



- [186] A. Turing. ‘Computing machinery and intelligence’. In: *Mind* 59 (1950), pp. 433–460 (cit. on p. 7).
- [187] F. Varela, E. Thompson and E. Rosch. *The embodied mind : Cognitive science and human experience*. Cambridge, MA: MIT Press, 1991 (cit. on p. 7).
- [188] J. X. Wang, Z. Kurth-Nelson, D. Tirumala, H. Soyer, J. Z. Leibo, R. Munos, C. Blundell, D. Kumaran and M. Botvinick. ‘Learning to Reinforcement Learn’. In: *arXiv:1611.05763 [cs, stat]* (2017) (cit. on p. 51).
- [189] J. Weng, J. McClelland, A. Pentland, O. Sporns, I. Stockman, M. Sur and E. Thelen. ‘Autonomous mental development by robots and animals’. In: *Science* 291 (2001), pp. 599–600 (cit. on pp. 5, 7).
- [190] Z. Xiao, X. Yuan, Q. V. Liao, R. Abdelghani and P.-Y. Oudeyer. ‘Supporting Qualitative Analysis with Large Language Models: Combining Codebook with GPT-3 for Deductive Coding’. In: *IUI 2023 - 28th International Conference on Intelligent User Interfaces*. Sydney, Australia: ACM, Mar. 2023, pp. 75–78. DOI: [10.1145/3581754.3584136](https://doi.org/10.1145/3581754.3584136). URL: <https://inria.hal.science/hal-04369097> (cit. on pp. 13, 61).