SPRIN–D

# Application for Challenge:

# Scientist-AI: An Adaptive Experimentation Platform

## Submitter


| | |
|---|---|
| Salutation: | Mr. |
| Title: | Dr. |
| Last name: | MAILLARD |
| First name: | ODALRIC-AMBRYM |
| Email: | ODALRICAMBRYM.MAILLARD@INRIA.FR |
| Phone number: | 0033643717048 |
| Your position: | CHARGE DE RECHERCHE INRIA, HABILITE A DIRIGER LES RECHERCHE, QUALIFIE PR SECTION 26 |
| Institution: | No Institution involved |
| Institution (other): | At the moment, no official institution involvement, though hosting support from Inria Studio is being discussed. |
| Legal Form: | Keine |
| Legal Form (other): | According to SPRIND requirements, appropriate company structure will be created. |
| Address line 1: | 172 Av. de Bretagne |
| Postal Code: | 59000 |
| City: | LILLE |
| Country: | France |


| | |
|---|---|
| Did you know about SPRIND...: | No |
| Where did you first learn...: | Recommendation by a friend / colleague |
| Please specify where you ...: | Former colleague from Inria |
| What made you decide to a...: | Colleague or collaborator encouraged me |
| Please specify why you de...: | This challenge uniquely enables a frontier AI effort combining academic RL research, full-stack engineering, scientific deployment, and entrepreneurship. Traditional |

academic funding rarely supports such long-term interdisciplinary operational structures.

## Core Idea

Scientist-AI (Scientist-ai.github.io) transforms AI from passive analysis of existing data into active optimization of experimentation. Combining reinforcement learning, uncertainty-aware decision-making, and domain-specific scientific platforms, it enables researchers to jointly formalize, simulate, and deploy adaptive experiments. The resulting infrastructure supports faster, safer, and more data-efficient discovery across precision medicine, agroecology, biology, and other experimental sciences.

## Description

### FRONTIER DIMENSION

#### Which “new frontier” does your idea target? E.g. model architecture (Mixture-of-Experts, State-Space-Models), data/training paradigm (RL-Verifiable-Reward, LoRA adapters), multi-agent orchestration, safety/alignment, or domain-specific AI application.

Scientist AI targets the frontier of Decisional Epistemic AI: a new training paradigm in which agents actively generate knowledge through real-world adaptive experimentation, contrasting with transformer-based in-silico content generation. Combining regret-minimizing Reinforcement Learning, world models, and digital twins, the system foster active information probing of biological, medical, agricultural, and ecological environments, closing Europe's strategic gap in AI with genuine epistemic autonomy

### CORE IDEA & ARCHITECTURE

#### Briefly describe the end-to-end system: main components, data flow, and how they interact.

While current frontier AI labs focus on LLMs and generative agents that synthesize text or scale massive in-silico datasets, they completely ignore the reality of physical experimental sciences. By contrast, real-world discovery happens in specialized, high-stakes environments where experiments are costly, slow, and constrained.

Scientist-AI explicitly targets this overlooked frontier: focused, medium-dimensional optimization problems (fewer than 10 dimensions) operating under strict, short sampling horizons (100 to 10,000 maximum samples).

Operating in this "small-data" regime is highly challenging: the system must be extremely data-efficient, inherently reproducible, safe, and entirely trustable by human practitioners. LLMs are structurally incapable of providing these mathematical guarantees. Our architecture bridges this gap by replacing black-box heuristics with structurally sound Bandit optimization and low-dimensional Markov Decision Process (MDP) learners.

Moreover, scientific AI needs domain-specific platforms for fields practitioners. Rather than simple UI wrappers, these are collaborative formalization interfaces built to bridge the translation gap between domain experts and ML researchers: They are intended to dramatically shorten the pre-experimental design phase (which traditionally takes months or years of manual alignment) by allowing practitioners to iteratively formalize objectives, constraints, and reward structures before a single physical trial begins.

A formalized problem maps onto the Adaptive Experimentation Toolkit to deploy data-efficient learners in real-world settings.
The system then executes a continuous Recursive Epistemic Loop:
- Observation: Modeling stochastic feedback from the experiment.
- Selection: Computing the next optimal intervention to maximize information gain or utility.
- Physical Execution: Running real-world experiments under strict operational constraints.

When domain simulators based on mechanistic models are available, the platform may leverage them for rapid prototyping and risk-free stress testing, serving as a powerful trust-building mechanism to convince the practitioner that the AI algorithm will behave safely and optimally before deploying it. Such mechanistic models are widely used in experimental sciences (e.g. Monolix for PK/PD modelling, DSSAT in agroecology, Pulse2Percept for retinal and cortical visual prostheses), and our platform could seamlessly interoperate with them. Availability and interoperability of mechanistic simulation frameworks is however not needed as the system is designed to operate directly in the real-world when no reliable simulator exists. Safety can be ensured by defining bounds on the space of possible actions prior to starting experiments.

The resulting architecture seamlessly unifies a structured Epistemic Engine, an Adaptive Decision Core, and collaborative domain platforms to shift AI from passive synthesis to active, real-world scientific discovery.

## TECHNICAL NOVELTY

### Which cutting-edge techniques (cite the most recent papers or pre-prints) are you combining, and why do they give a measurable advantage over the current state-of-the-art? Which capability gap in current AI systems, unlikely to be addressed through the continued development of existing systems, does this submission address?

We propose a novel integration of regret minimization Reinforcement Learning (RL) specifically tailored for the experimental sciences, combining the Confusing Instance (CI) Principle with World Models and taylored-domain Digital Twins.

We combine the following most recent approaches:

- The Confusing Instance (CI) Principle: Unlike Optimistic or Bayesian approaches, the CI principle directly leverages the fundamental performance limits for a specific problem instance, constructing statistically similar alternative models that prescribe different optimal behaviors. This approach allows for dramatic improvements in sample efficiency, as detailed in hal-03825423, arxiv.org/

pdf/2501.13013, arxiv.org/pdf/2510.19531, arxiv.org/pdf/2510.21232, including in continuous and weakly communicating environments.

- Bandit optimization for various experimentation constraints, organized in a coherent stack: e.g. Non-parametric (hal-03421252) or corrupted rewards (hal-04615733), risk-aversion (hal-03447244), or context-drift (openreview.net/forum?id=qJZ9EEmff5), as developed by the marlelgroup.github.io/ group.

- Digital Twin Integration: specialized world models (e.g. gym-dssat for agroecology, STED microscopy models for super-resolution imaging), and human-in-the-loop optimization (HILO) to personalize retinal and cortical implants (https://arxiv.org/abs/2306.13104)

We address a major gap in Agentic-AI: current systems often invoke external tools based on passive knowledge bases, but fail to leverage active experimentation algorithms. We develop a bandit/MDP-based toolkit that enables Agentic-AI to search and use these tools in a statistically sound and coherent way. We can further leverage dynamics (e.g., from a Digital Twin) to optimize exploration. This transforms AI from a generative tool into a Decisional Agent capable of iterative hypothesis testing and optimization in uncertain, high-dimensional scientific environments, assisting scientists in knowledge discovery.

## TECHNICAL NOVELTY CITATION

Bandit-optimisation feasibility (short/moderate horizon):
[Bio] https://www.nature.com/articles/s41467-018-07668-y.pdf
[Agro] https://arxiv.org/pdf/2210.04537
Marlel group publications https://marlelgroup.github.io/

Confusing instance paradigm:
https://inria.hal.science/hal-03825423/
https://proceedings.mlr.press/v222/saber24a/saber24a.pdf
https://arxiv.org/pdf/2510.19531
https://arxiv.org/pdf/2510.21232
https://arxiv.org/pdf/2501.13013

Bandits Variants (Risk-averse, Non-parametric, Corrupted, Batch, Context-drift):
https://inria.hal.science/hal-03447244/
https://hal.science/hal-04615733/
https://inria.hal.science/hal-03421252/
https://openreview.net/forum?id=qJZ9EEmff5

Human-in-the-loop optimization in medicine:
https://proceedings.neurips.cc/paper_files/paper/2023/file/fb06bc3abcece7b8725a8b83b8fa3632-Paper-Conference.pdf
https://iopscience.iop.org/article/10.1088/1741-2552/ac7615
https://arxiv.org/pdf/2502.00177
https://drive.proton.me/urls/X9YSBZWBGG#ABmSyexNoidm

## CAPABILITY GAP ADDRESSED

### Explain how the solution creates new functionality (efficiency, reasoning, controllability, new use-cases) rather than just copying existing transformer models.

While transformers passively synthesize existing literature, Scientist-AI introduces Epistemic Autonomy: the capability to actively probe unknown environments, select sequential interventions to generate novel scientific data. We target four paradigm-shifting functionalities:

- Controllability & Soundness: Grounded in regret-minimization theory, the platform provides rigorous mathematical guarantees on exploration risk, bypassing often unreliable deep RL heuristics.
- Extreme Data Efficiency: Rather than requiring millions of simulated trajectories, our algorithms optimize for short-horizon and limited trajectory experimental constraints.
- Human-in-the-Loop Reasoning: Contrasting generic black-box inference, we assist practitioners in co-specifying objectives, safety bounds, and domain semantics, keeping humans in absolute control.
- New Critical Use-Cases: We unlock high-stakes, real-world deployment in safety-critical, data-scarce sectors like adaptive precision medicine and agroecology.

## EXISTING ARTIFACTS

### What prototypes, code bases, data pipelines, or models are already working? Attach links or short demos if possible.

Several operational building blocks already underpin Scientist-AI.

Algorithmic core. The StatisticalRL ecosystem (github.com/StatisticalRL) provides research-grade implementations of Bandit and MDP algorithms under uncertainty, delayed feedback, and constrained settings, supporting reproducible benchmarking and rapid prototyping. This is enriched with MARLEL publication codes.

Agroecology pilots: Adaptive Design-of-Experiments strategies were successfully deployed for crop management optimization (arxiv.org/pdf/2210.04537). gym-DSSAT (rgautron.gitlabpages.inria.fr/gym-dssat-docs) offers a Gymnasium-compatible mechanistic simulator interface; additional simulators (WOFOST-gym, STICS, MAELIA) and European climate datasets are available (hal-03231805).
European datasets on agroecological experimentations are available https://zenodo.org/records/14160885
from our Inrae collaborators.

Biology & medicine: Simulation-assisted adaptive microscopy workflows demonstrated real-world experimentation loops (nature.com/articles/s41467-018-07668-y). A patent application (FR 2510219) covers a human-in-the-loop Levodopa dosing optimizer combining PK/PD modelling and Bayesian optimization for Parkinson's patients. This is completed by progress in human-in-the-loop optimization for retinal implants ieeexplore.ieee.org/document/11253762

A biology formalization mockup showing potential multidiscplinary interface is live at ai.studio/

apps/44d952c0-267e-4ad8-bfa2-687196f494fe.

Field deployments: Tisanji deploys bandit-based decision wrappers in horticulture classrooms across 4 schools (tisanji.com/bandit). ComTech Alliance operates SautiRafiki (multilingual voice AI), SHAURI (agricultural advisory), and SHIRIKI (farmer peer-learning) in Sub-Saharan Africa.
Both contribute to Knowlege Equity.

Community: The MARLEL group is organizing a dedicated RL-for-experimental-sciences workshop in December (rl-experimental-sciences-workshop.github.io), instrumental for Phase II ecosystem building.

## TECHNOLOGY READINESS LEVEL (TRL) ASSESSMENT

### Rate the maturity of your overall system and each major sub-component.

The overall Scientist AI vision stands between TRL 2 and TRL 4. Core adaptive decision-making algorithms are validated in constrained settings. The integrated multidisciplinary platform architecture remains at an earlier prototyping stage.

Layer 1 - Adaptive Experimentation Toolkit (TRL 4–5):
Pilots - e.g. gym-DSSAT – and algorithmic building blocks - e.g. bandits, RL, constrained optimization, delayed feedback, simulation interfaces - provide realistic Gymnasium-compatible environments for experimentation research.

Layer 2 - Domain-Specific Scientific Platforms (TRL 2–3):
Proof-of-concepts involving ML researchers and domain scientists, support initial experimentation specification, reward prototyping, and simulator integration, but require substantial engineering and scaling.

Layer 3 - Deployment and Real-World Integration (TRL 3–4):
Prototypes in biology, agronomy, and industrial systems have validated feasibility on real scientific workflows. Large-scale deployment remains a Stage 2–3 objective.

## OPEN RESEARCH RISKS

### Identify the parts that are still research-grade.

Beyond established foundations it remains:

1] Transferring theoretically grounded decision-making methods to real scientific workflows combining uncertainty, delayed observations, safety requirements, limited budgets, and heterogeneous constraints. We mitigate this by focusing initially on a small number of high-value use cases in agroecology and precision medicine.

2] Sim-to-real transfer. Digital twins and simulators only partially capture biological, agricultural, or clinical systems. We address this through simulation-first validation, uncertainty-aware decision-making, and progressive deployment under human supervision.

3] Adoption & formalization. Translating scientific questions into executable adaptive experimentation workflows remains largely manual. Dedicated domain platforms co-designed with practitioners directly target this bottleneck. We propritize enforcing ethical constraints, knowledge equity and ensuring human experts retain decision authority at any execution step.

## COMPUTE REQUIREMENTS

### List compute required for Stage 1, include total hours needed, GPU type, source and justification.

We will prioritize acquiring CPU usage, and small local GPU ressources, with only occasionally access to large GPU resources, given the focus on phase I away from LLMs.
- 128-core CPU (e.g., AMD EPYC 7643, Intel Xeon)
  - Simulator + RL/bandit research are mostly CPU-bound, 128 cores = 128 parallel simulations
  - RAM: 128-256 GB
- 2–4× Mid-range GPU (e.g., RTX 4090, 5080, 5090)
  - Deep RL policy training, (DQN, PPO, A3C..)
  - GPU based environments,
  - hyperparameter search and statistical validation
- Cloud GPU Rental (e.g., A100, L40S). Credits for 300 hrs/mo × 6 = 1,800 GPU-hours
  - Fine-tuning domain LLMs for hypothesis generation
- Storage + MLOps Infrastructure
  - High-speed NVMe: 2-4 TB
  - S3-compatible object storage: 5-10 TB
  - Experiment tracking (W&B, DVC, MLflow)

The academic nature of the project enables applications for access to French research computing infrastructures such as Jean Zay Supercomputer and Grid'5000.

## KPIS, BENCHMARKS AND POTENTIAL IMPACT

### Which KPIs or benchmarks (if existing) would you use to track your progress? In what existing or future industry do you expect the highest impact? How would you assess it?

Stage I focuses on validating the operational hypothesis underlying Scientist AI: that experimental scientists and decision-making experts can jointly formalize adaptive experimentation problems, connect them to modern RL & sequential optimization algorithms, and obtain reproducible end-to-end experimentation pipelines.
Success will be evaluated through algorithmic, scientific, and operational KPIs. Algorithmic KPIs include safety-constraint violations, robustness to delayed/corrupted feedback, sample and computational efficiency on RL/Bandit benchmarks and domain simulators. Scientific KPIs include reductions in failed experiments, experimental budgets, and trials required to reach target outcomes. Operational KPIs include multidiscipinary formalization time, practitioner compliance with recommended actions, completion of demonstrators in agroecology & precision medicine, and engagement of pilot users and partners. Broader discovery cycles and economic impacts will be assessed in later stages.

WORK PLAN

The plan should:
• Cover milestones for the entire duration of the three stages, with specific information on what will be achieved after Stage 1.
• Indicate where collaboration and subcontracting work packages will be required.
• Contain cost information, including contributions by the team or the team's institution (e.g. personnel, licenses, infrastructure) for three stages of the Challenge.

Stage 1 - Operational Foundations and Prototype Validation

Stage 1 is focused on rapid operationalization rather than broad conceptual expansion. It aims to validate the core Scientist AI hypotheses via working prototypes, real scientific use-cases, and scalable organizational foundations.

WP1 concerns development of the Experimental Science RL Toolkit, a modular adaptive experimentation library inspired by the simplicity and interoperability of scikit-learn. Building on existing artifacts such as the StatisticalRL package, the toolkit will integrate state-of-the-art RL/Bandit algorithms, uncertainty-aware optimization, batch experimentation, safety constraints, and simulator interfaces. The objective is to provide reusable APIs enabling rapid adaptive experimentation prototyping across scientific domains.

WP2 focuses on domain-specific experimentation platforms for Agroecology and Precision Medicine. Rather than abstract interfaces disconnected from practice, the project will rely on real experimentation workflows proposed by partners such as ComTech Alliance and Tisanji. The platforms will facilitate multidisciplinary dialogue between ML researchers and practitioners through collaborative formalization tools, reward shaping interfaces, simulator integration, and constraint specification. Initial user engagement and iterative feedback loops will already be established during Stage 1.

WP3 concerns infrastructure and operational scalability: deployment pipelines, authentication, benchmarking, experiment tracking, interoperability between platforms and toolkit, and preparation of the company structure through the Inria Startup Studio incubator.

WP4 addresses frontier RL research on adaptive experimentation, regret-optimal learning, short-horizon RL, uncertainty-aware decision-making, batch-constrained experimentation, and the Confusing Instance paradigm. Benchmark comparisons will be conducted on synthetic environments and domain-specific simulators.

Subcontracting and collaborations are central to the project.
E.g., ComTech Alliance will contribute agroecology deployment workflows, field interfaces and experimentation constraints. Additional scientific workshops and invited missions will strengthen connections with MARLEL and European RL ecosystems.

Stage 1 deliverables include:

* Open-source prototype of Experimental Science RL Toolkit
* Agroecology and Precision Medicine demonstrators
* Benchmark suite and evaluation protocols
* Scientific report validating frontier hypotheses
* Adjusted roadmap for Stage 2
* Company structure and operational R&D organization

The estimated Stage 1 budget is about €1.5M–€2M, primarily dedicated to engineering personnel, infrastructure, scientific coordination and strategic partnerships.

Stage 2 - Scaling and Stress Testing

Stage 2 focuses on scaling the platform technically and operationally. The toolkit will be extended to larger state/action spaces, hybrid architectures, multi-agent settings, and higher-dimensional experimentation. Domain platforms will integrate additional simulators, real data pipelines, and broader practitioner communities.

Stress tests will include delayed or corrupted feedback, higher data volumes, larger agent populations, sim-to-real transfer, and heterogeneous operational constraints. Additional scientific domains may be progressively incorporated.

Stage 3 - Deployment and European Adaptive Experimentation Infrastructure

Stage 3 will transform Scientist AI into a scalable European adaptive experimentation infrastructure. This phase includes large-scale deployment pipelines, enterprise-grade infrastructure, governance and safety frameworks, sustainability and business development, and integration into European academic and industrial ecosystems.

The long-term objective is to establish Scientist AI as a sovereign European platform bridging frontier RL research and real-world scientific discovery.

## FINANCIAL COST ESTIMATE FOR STAGE 1

### Max 3,000,000 EUR, in EUR.

1,633,049 EUR

## TEAM

### Address:

### Why is your team best suited to achieve the goal of the challenge?

### What relevant field-specific expertise and experience do you already have, and with whom?

### What expertise is the team missing, and how will you address it?

### Attach the CVs of the key people in the Attachments section.

Scientist AI unites frontier RL research, scientific experimentation expertise, and large-scale engineering in a combination rarely found in a single team, precisely what is needed to turn sequential decision-making advances into operational tools for experimental scientists.

Scientific direction is led by Odalric-Ambrym Maillard (Inria, MARLEL, ELLIS), an internationally recognised expert in RL, bandits, and adaptive experimentation, with a decade of interdisciplinary collaborations in agronomy and medicine. He is complemented by Tristan Fauvel (PhD computational neuroscience), whose track record spans clinical trials and human-in-the-loop Bayesian optimisation and patent in precision medecine application. Three senior scientific advisors (Ortner, Jonsson, Talebi) cover regret-optimal MDPs, hierarchical planning, and risk-sensitive uncertainty quantification. Engineering leadership is provided by Pierre Gronlier (former CTO of Gaia-X, federated cloud at European scale) and Umberto Colella (founder-CTO, critical infrastructure, reproducible deployment).

Domain grounding comes from Krishna Naudin (CIRAD, 25+ years agroecology), Suzanne Reynders (INRAE, EU Carbon Credit framework), and Devotha Nyambo (ComTech Alliance), who brings field-tested RL and LLM deployments in Sub-Saharan Africa and ensures knowledge equity. Tisanji's live bandit platform, already used in Canadian horticulture schools, validates the pedagogical deployment pathway.

The principal gap is product and business leadership. This will be filled through the recruitment of a dedicated CEO, supported by the Inria Startup Studio incubator and Raphael Attias (European R&D programs, UN advisor).

Last, our team has good ties with generative AI companies outside the US (Mistral, Cohere): Though not required for the core project, future incorporation of diverse language-based interactions may benefit from targeted collaborations with these actors.

## TEAM MEMBER 1

### Name, role on the project and percent FTE committed

Odalric-Ambrym Maillard, CR HDR, Principal Research Scientist, 40%FTE

## TEAM MEMBER 1

### Links to demonstrate track record (Industry, Entrepreneurship, Google-Scholar, GitHub-Project)

https://odalric-ambrymmaillard.github.io/researcher/publications.html
https://marlelgroup.github.io/
https://odalric-ambrymmaillard.github.io/researcher/research_projects.html
https://github.com/StatisticalRL
https://github.com/farm-gym

TEAM MEMBER 2

### Name, role on the project and percent FTE committed

Tristan Fauvel, Machine Learning Engineer
100% FTE

TEAM MEMBER 2

### Links to demonstrate track record (Industry, Entrepreneurship, Google-Scholar, GitHub-Project)

https://scholar.google.com/citations?user=kJ10450AAAAJ
Patent application n°FR 2510219, "DEVICE FOR GENERATING PATIENT-SPECIFIC DRUG ADMINISTRATION PATTERNS".
2026 World Parkinson Congress:
https://drive.proton.me/urls/VEXRMQ04RC#JDAltyWeKu4Z

TEAM MEMBER 3

### Name, role on the project and percent FTE committed

We are accompanied by CEO hiring expert, Raphael Attias, specialist in European R&D programs to help find Team Member 3

## Declarations

I/we declare that I/we agree to the processing of my/our data in accordance with the SPRIND GmbH (Next Frontier AI Challenge) privacy notice - dated April 2026.

I/we agree with the statement on reasons for exclusion:: Yes

I/we agree to require the above declaration also from subcontractors and to submit it to SPRIND upon request.
: Yes

## Meta Data

- Form Type: challenge
- Form Version: 0.2
- Language: EN
- Date: 01.06.2026 09:55
- Number of Documents Attached: 4
- Checksum: e5b6841880d876535766df97966eaf4703a5ada5