

# Bryan Lincoln Marques de Oliveira

AI RESEARCH SCIENTIST · PHD STUDENT IN COMPUTER SCIENCE · CEIA / AKCIT / UFG

✉ bryanlmoliveira@gmail.com | 🏠 bryanoliveira.github.io | 📧 bryanoliveira | 📄 bryanoliveira | 🐦 @bryanlmoliveira | 🎓 Scholar

## Summary

AI researcher specializing in Deep Reinforcement Learning. PhD Student at UFG and leading interdisciplinary teams at CEIA/AKCIT on reinforcement learning for decision-making systems, emphasizing agency, reasoning, and information-seeking behavior. Advancing research towards generalist agents that learn from experience while bridging theoretical foundations with large-scale deployment.

## Research Interests

- **Deep Reinforcement Learning** — Policy optimization, representation learning, credit assignment, offline RL, real-world deployment
- **Decision Making Under Uncertainty** — Information theory, Bayesian inference, uncertainty quantification, information-seeking behavior
- **Robotics & Embodied AI** — VLAs, teleoperation, sim-to-real transfer, digital twins
- **Agent Systems** — Multi-agent harnesses, embodied planning, lifelong learning
- **LLM Post-Training & Alignment** — RLHF, RLAI, RLVR, reasoning, tool use, interpretability, and steerability

## Education

### Universidade Federal de Goiás

Goiânia, Brazil

PHD IN COMPUTER SCIENCE

2026 – 2030

- Advisor: Prof. Dr. Telma Woerle de Lima Soares
- Focus: Decision Making Under Uncertainty, Reinforcement Learning, Agent Systems

### Universidade Federal de Goiás

Goiânia, Brazil

M.Sc. IN COMPUTER SCIENCE

2022 – 2025

- Advisor: Prof. Dr. Telma Woerle de Lima Soares
- Thesis: *Sliding Puzzles Gym: A Scalable Benchmark for State Representation in Visual Reinforcement Learning* (published at ICML 2025)
- Proposed SPGym — a scalable extension of the 15-tile puzzle to evaluate RL algorithms, revealing opportunities for advancing representation learning for decision-making.
- Keywords: Reinforcement Learning · Representation Learning · Benchmark · Visual RL

### Universidade Federal de Goiás

Goiânia, Brazil

B.Sc. IN COMPUTER SCIENCE

2015 – 2019

- Advisor: Prof. Dr. Telma Woerle de Lima Soares
- Thesis: *Intrinsic Motivation for Robotic Manipulation Learning with Sparse Rewards*
- Keywords: Reinforcement Learning · Intrinsic Motivation · Sparse Rewards · Robotics

## Selected Experience

### AKCIT — Advanced Knowledge Center for Immersive Technologies

Goiânia, Brazil

STAFF RESEARCH SCIENTIST

Nov. 2024 – Present

- Leading basic research and PoC development across 5 sub-teams:
- **Information-Seeking Agents:** Dialogue agents for optimal information-seeking under uncertainty via information theory (arXiv:2601.17716).
- **Policy Gradient Methods:** Theoretical and empirical study of GRPO and PPO in classical RL environments (arXiv:2511.03527).
- **Embodied Agents:** LLM-based embodied planning under uncertainty, Bayesian belief-state tracking.
- **Autonomous Humanoids:** Teleoperation, policy training, digital twins, Isaac Sim, sim-to-real transfer.
- **Small Language Models:** Continued pre-training and post-training RL for offline Brazilian Portuguese LLMs in education.
- Mentoring junior researchers across all sub-teams.

### CEIA — Centro de Excelência em Inteligência Artificial

Goiânia, Brazil

STAFF RESEARCH SCIENTIST & ML ENGINEER

Aug. 2019 – Present

- R&D from PoC to production-grade deployment across finance, energy, tourism, and agribusiness.
- **LLM Agents for Business** (in partnership w/ HUG Labs, Nov 2024–present): Conv-to-Bench, a multi-turn LLM evaluation benchmark from raw interaction logs (arXiv:2603.13626); Deep research agents for business documents; RLM-based agent that implements and evolves other agents.
- **LLM Post-Training** (in partnership w/ CEMIG, Dec 2023–present): RLHF, RLAI, and RLVR pipelines for alignment, reasoning and tool-use capabilities in a large language model for the Brazilian energy sector.
- **Debt Pricing & Collection** (in partnership w/ Recovery/Itaú): Applied RL to orchestrate debt pricing strategies (RLC 2025 Workshop).
- **Marketplace Personalization (in partnership w/ Bancorbras):** RL for recommendation systems for travel offers. (RLC 2025 Workshop).
- **Advertising Optimization** (in partnership w/ bettrAds): RL and generative models for multi-channel ad generation and campaign optimization.
- **Offline RL for Digital Marketing:** Conservative Q-Learning for personalized marketing (NeurIPS 2021 Offline RL Workshop Oral).
- Automated scalable ML pipelines (Luigi, BentoML, Seldon, FluxCD, Kubernetes); terabyte-scale ETL (BigQuery, Spark).

## Data H Artificial Intelligence

Remote

ML ENGINEER / FULL STACK SOFTWARE ENGINEER

Feb. 2020 – Jan. 2022

- Built anti-plagiarism system for LMS platforms (Facenet, DeepSpeaker, BERT, ElasticSearch) and energy forecasting & anomaly detection system (Prophet, PyTorch Forecasting, Optuna).
- Deployed scalable ML pipelines (Luigi, DataProc, BentoML, Seldon, Kubernetes); full-stack development (Django, VueJS, PostgreSQL).

## MusicPlayce

Goiânia, Brazil

LEAD BACKEND & INFRASTRUCTURE ENGINEER

Jul. 2019 – Feb. 2020

- Led scalable back-end development (Python, MongoDB, Neo4J); cloud microservice architecture on GCP/AWS with GitLab CI/CD.

## Publications

---

### Conv-To-Bench: Evaluating Language Models via User-Assistant Dialogues in Code Tasks

ICLR 2026 — DATA-FM Workshop

VM DOS SANTOS, AC CASTRO, SLS TOLEDO, BML CALURA, LCM MENEZES, RCR MATA, TWL SOARES, **BRYAN LM DE OLIVEIRA**

2026

### Do Reasoning Models Ask Better Questions? A Formal Information-Theoretic Analysis on Multi-Turn LLM Games

AAAI 2026 — NeusymBridge Workshop

DM PEDROZO, TWL SOARES, **BRYAN LM DE OLIVEIRA**

2026

### Partial Reasoning in Language Models: Search and Refinement Guided by Uncertainty

AAAI 2026 — LaCATODA Workshop

ML DA LUZ, B BRANDÃO, LGB MARTINS, G OLIVEIRA, **BRYAN LM DE OLIVEIRA**, LC MELO, TWL SOARES

2026

### Sliding Puzzles Gym: A Scalable Benchmark for State Representation in Visual Reinforcement Learning

ICML 2025

**BRYAN LM DE OLIVEIRA**, LGB. MARTINS, B BRANDÃO, ML LUZ, TWL SOARES, LC MELO

2025

### Learning Without Critics? Revisiting GRPO in Classical Reinforcement Learning Environments

NeurIPS 2025 — LatinX in AI Workshop

**BRYAN LM DE OLIVEIRA**<sup>\*</sup>, FV FRUJERI<sup>\*</sup>, MPCM QUEIROZ<sup>\*</sup>, LGB MARTINS, TWL SOARES, LC MELO.

2025

### InfoQuest: Evaluating Multi-Turn Dialogue Agents for Open-Ended Conversations with Hidden Context

RLC 2025 — RLBrew Workshop (Oral)

**BRYAN LM DE OLIVEIRA**, LGB MARTINS, B BRANDÃO, LC MELO

2025

### Reinforcement Learning for Debt Pricing: A Case Study in Financial Services

RLC 2025 — RL for Real-World Systems Workshop

B BRANDÃO<sup>\*</sup>, LGB MARTINS<sup>\*</sup>, **BRYAN LM DE OLIVEIRA**<sup>\*</sup>, LC MELO, ML LUZ, EAS GARCIA, MV DA SILVA, RG AVELAR, ARG FILHO, AS SOARES, TWL SOARES

2025

### Personalizing Fairness: Adaptive RL with User Diversity Preference for Recommender Systems

RLC 2025 — RL for Real-World Systems Workshop

LGB MARTINS, **BRYAN LM DE OLIVEIRA**, B BRANDÃO, TWL SOARES, MRO SANTANA

2025

### PulseRL: Enabling Offline Reinforcement Learning for Digital Marketing Systems via Conservative Q-Learning

NeurIPS 2021 — Offline RL Workshop (Oral)

LC MELO<sup>\*</sup>, LGB MARTINS<sup>\*</sup>, **BRYAN LM DE OLIVEIRA**<sup>\*</sup>, B BRANDÃO<sup>\*</sup>, DWR SOARES, TWL SOARES

2021

## Technical Skills

---

- **Reinforcement Learning:** Policy gradient methods (GRPO, PPO, DPO, Actor-Critic), Credit assignment, Offline RL (CQL, behavior cloning), Model-based RL, Curiosity & intrinsic motivation
- **Decision Making & Probabilistic ML:** Bayesian inference, Information theory, Uncertainty quantification, Bandit algorithms, Sequential decision making
- **Robotics & Simulation:** MuJoCo, Isaac Sim, Sim-to-real transfer, Motion capture retargeting, Teleoperation systems, OpenCV
- **LLM Post-Training:** RLHF, RLAI, RLVR, SFT, DPO, LoRA/QLoRA, Reward modeling, Alignment
- **Deep Learning / ML:** PyTorch, TRL, Hugging Face Transformers, Representation learning, Computer vision, Recommendation systems
- **Infrastructure / MLOps:** SLURM/HPC, Kubernetes, Docker, Ray RLLib, vLLM, Ollama, BentoML, BigQuery, Spark
- **Research Leadership:** Multi-team management, Junior researcher mentoring, R&D project scoping and execution, Cross-institutional collaboration, Scientific writing and communication
- **Languages:** Portuguese (Native) · English (C2 Proficient, EF SET 77/100)

## Extracurricular Activity

---

### Institut d'Intelligence Artificielle Appliquée (I2A2)

*Goiânia, Brazil*

STUDENT TRAINEE

*Jul. 2019 – Feb. 2020*

- Implemented and trained CNN models like ResNet using PyTorch to perform bone age regression.
- Studied recurrent neural networks and transformers. Applied BERT to perform Named Entity Recognition tasks on legal texts.
- Implemented LDA and TF\*IDF algorithms to learn about topic modeling methods.

### Deep Learning Brasil

*Goiânia, Brazil*

UNDERGRADUATE RESEARCH SCHOLAR

*Mar. 2019 – Jul. 2020*

- Studied and implemented RL and Deep RL algorithms like DQN and PPO, applying them to continuous control tasks.
- Experimented with curiosity and intrinsic motivation methods in sparse-reward RL environments.
- Studied, implemented, and experimented with curiosity-like techniques applied to meta-RL algorithms.
- Studied autoencoders, variational inference, PCA, and embedding techniques to support meta-RL experiments.

### Núcleo de Robótica Pequena Mecânica

*Goiânia, Brazil*

TEAM LEADER & C++ DEVELOPER

*Feb. 2017 – Nov. 2018*

- Coordinated a team of 6+ interdisciplinary developers using Scrum; led development of the robot control system (C++, Python, OpenCV, Qt). System later used as benchmark in a multi-agent deep RL Master's degree project.
- Developed computer vision systems for real-time robot tracking; strategy and navigation algorithms using vector fields; GTK interfaces for real-time monitoring.
- Competed in LARC 2017 (4th place, VSSS) and LARC 2018; organized Arduino/robotics workshops for the community.
- Developed MuJoCo-based integrated simulator for fast navigation and strategy experiments.

### Federal University of Goiás

*Goiânia, Brazil*

TEACHING ASSISTANT OF COMPUTER PROGRAMMING / DISCRETE MATHEMATICS

*Aug. 2015 – Dec. 2017*

- Assisted students with data structures, algorithms, logic, formal proofs, and C/C++ concepts. Supported teachers with assignment elaboration and grading.