



Michal Valko

Founding Researcher · LLMs · RLHF · Self-Supervised Learning · Bandits



[misovalko.github.io](https://github.com/misovalko)

michal.valko@inria.fr

159

Papers

11k+

BYOL Cites

54

h-index

15

PhD Grads

10+

Grants



BEST PAPER

CAREER

- **Isara Labs** Founding Researcher 2025–present
- **Meta · GenAI** Principal, Llama 3 2024
- **DeepMind** Sr. Staff Scientist 2019–2024
- **ENS Paris-Saclay** MVA Lecturer 2014–present
- **Inria Lille** Tenured Researcher 2012–present
- **Intel Labs** FacelD AI Researcher 2009–2010

EDUCATION

- **ENS Cachan HdR** Mathematics 2016
- **U. Pittsburgh PhD** ML, GPA 4.0 2011
- **CMU** Visiting Student, CS & Stats 2007–2009
- **Comenius U. MSc** Math & CS 2005
- **Gym. Alejová, Košice** High School 2000

RESEARCH FOCUS

LLM Alignment RLHF Reasoning

Fine-tuning World Models Bandits

RL Theory Self-Supervised Graphs

Exploration DPPs Kernels

RAG MCTS GNNs Planning

GRANTS & FUNDING

- 8 grants as PI/coordinator, €12M+ raised
- BOLD (ANR) 19–23, Coordinator
 - Extra-Learn (ANR) 14–18, PI
 - DELTA (EU) 18–22
 - CompLACS (EU FP7) 11–15
 - EduBand with CMU, IoT (PI), NIH

SERVICE

- Inria: Eval (2x), Hiring (3x) Comm.
- **AE:** TMLR, JMLR, MLJ; **AC:** NeurIPS, ICLR, IJCAI
- **PC:** COLT, ICML, AISTATS, ALT, UAI, KDD
- Top Rev: ICML '15, '18; Grant: FNRS, ISF
- Workshop: EWRL, SSL, Bandits, RLSS
- EEML '23 co-chair; **Advisor:** Archemind
- 15+ PhD/HdR jury, 3 Habilitation comm.

RECOGNITION

- 🏆 ICML 2023 Outstanding Paper
- ★ Forbes Cover — Slovak Alnstein
- ★ Forbes Changemakers 2024 — AI
- 🏆 ASAI AI Impact #1; Innovator #2 '26
- ★ AI Summit (UN/OECD/UNESCO/ECB)
- ★ Inria Excellence 2x (14–17, 18–21)
- 🏛 Comenius Alumni '15
- 🎓 Mellon Fellowship '08–09
- 🏆 Homer Warner (AMIA) '10
- 🏆 Google Best Paper '10
- 🏆 Compunetix Award '08, '11
- 🏆 Slovak Academy Fellowship '03–05

SELECTED RESEARCH IMPACT

🦋 Llama 3

Meta 2024 · Principal Llama Scientist, building the online RL stack and research for Llama 3.

🦋 BYOL

DeepMind 2020 · Self-supervised learning without negatives. 11k+ citations. Foundation for modern SSL methods.

🦋 Nash-MD / NLHF

DeepMind 23–24 · Game-theoretic preference learning: alignment via Nash equilibrium rather than only scalar-reward maximization.

🦋 Spectral Bandits

ICML 2014 · Graph Laplacian bandits with regret guarantees. Kernel-based exploration in structured spaces.

🦋 DPPy Library

Open Source · Python library for Determinantal Point Processes. ★ 500+ GitHub stars. Widely used in ML.

🦋 World Models

DeepMind 20–24 · Curiosity-driven exploration and intrinsic motivation. Representation learning for long-term novelty.

🦋 SQUEAK

Inria 09–18 · Online sparsify kernels and graphs. Single-pass streaming dict learning with guarantees.

SELECTED PAPERS

- 📖 Llama 3 Herd of Models (2024)
- 📖 Nash Learning from Human Feedback
- 📖 Game Trees 🏆 ICML Best (poster, talk)
- 📖 BYOL (NeurIPS 20) (code)
- 📖 BYOL-Explore (ICML 22) (poster, talk)
- 📖 Distributed SQUEAK (AISTATS 17)
- 📖 Spectral Bandits (ICML 14) (slides, poster)
- 📖 Unified RLHF/DPO/IPO (AISTATS 24)

INVITED TALKS

- **Keynote:** COLT 2024, EEML 2023
- OECD AI 2025, WSJ, Forbes AI Summit
- DeepMind, EWRL 2023, RLSS

MEDIA & PRESS

- WSJ (26, 20), NY Post, BI, Les Echos, engadget, týždeň, Forbes, The Information
- Denník N, Radio Slovakia, Slovak Spectator, Vedátor, ŽIVÉ.SK

PHD STUDENTS & POSTDOCS

Current:

- G. Racca (U. Copenhagen, 25–28)
- A. Špitalský (Master, 25)
- A. Munos (Mentoring, 25)

Graduated:

- L. Bedin → Cartesia
- J.-B. Grill → DeepMind
- D. Calandriello → DeepMind
- D. Tiapkin → DeepMind
- J. Tarbouriech → DeepMind
- Y. Tang → Anthropic
- C. Fiegel → ENSAE
- O. D. Domingues → Cohere
- X. Shang → QRT
- J. Seznec → Lycée LdV
- P. Perrault → IDEMIA
- G. Gautier → Decathlon
- T. Kocák → U. Potsdam

Postdocs:

- P. Ménard → Meta
- É. Oyallon → CNRS

RESEARCH THEMES

LLM Alignment: Nash, IPO, Llama 3.

Zero-Sum Games: Imperfect info, 🏆 Best.

Self-Supervised: BYOL, 11k+ cites.

World Models: Curiosity, intrinsic.

Exploration: BYOL-Explore, reward-free.

Off-Policy RL: Value functions, credit.

Graph Bandits: Spectral, side info.

Kernel Methods: SQUEAK, online sparsif.

Graph Learning: DPPs, semi-supervised.

Sample Efficiency: MCTS, planning, bandits.

Online Learning: Adaptive alg., regret bounds.

TEACHING

- **Graphs in ML** — MVA, 2014–now
- 15 PhD, 2 Postdocs, 20+ MSc supervised
- 1500+ students taught, 10+ jury

STARTUPS & INVESTING

- VC: Sparkle, Formula, KAYA, Cherry, Visionaries, RAISE; Angel: Entalpic, Forgent
- Recent: Mistral AI, Mercor, Comand AI, Manifold
- 30+ startups advised incl. Holonomy, Sword, Doctolib, Agemo, Tortus, Tzafon, Atla, HinLab, Robi, Prefrontal AI

BEYOND RESEARCH

- 🎵 Coeli et Terra, 🧑🏃 Hiking, Squash
- 🎮 SK, EN, FR, CZ