

P é t e r J ó z s e f L ó d r i → Sakana AI (サカナAI)

Role: AI Systems & Multi-GPU Infrastructure Engineer (Low-Bit Inference & Agentic Swarms) · Comp:

\15,000,000 – \28,000,000+ JPY + Equity · Location: Minato-ku, Tokyo (港区・虎ノ門 / 赤坂)

Dedicated Subdomain: <https://sakana.vaked.dev> · Portal: <https://sakana.ai/careers/> (careers@sakana.ai)

01. COVER LETTER & OUTREACH MESSAGE (INMAIL / EMAIL)

Dear Sakana AI Founding & Engineering Team,

I built a dedicated interactive showcase and bilingual engineering dossier for Sakana AI's AI Systems & Multi-GPU Infrastructure Engineer role in Tokyo:

- Live Interactive Portfolio (Emergent Boids + Ternary Lattice): <https://sakana.vaked.dev>
- Tailored Resume / 職務経歴書 (PDF): <https://sakana.vaked.dev/resume.pdf>
- Systems Portfolio (PDF): <https://sakana.vaked.dev/portfolio.pdf>
- Full Application & Architecture Note: <https://sakana.vaked.dev/application>

My work centers on two pillars at the heart of Sakana AI's mission:

1. Low-Bit Foundation Model Compression (8b-is/MLX-QUANT & pureQTern): 12.80× memory reduction via BitNet b1.58 {-1, 0, +1} ternary quantization with custom Metal GPU kernels and zero-dependency C/Rust codecs.
2. Collective Multi-Agent Systems: Core Architect on Osaurus (osaurus-ai/osaurus — native Swift offline agent harness), contributor to Headroom (headroomlabs-ai/headroom PR #1418 — kompress domain-routing bias multipliers), and builder of AgentField multi-reasoner DAG orchestration.

Warm regards,

P é t e r J ó z s e f L ó d r i

<https://www.linkedin.com/in/peter-lodri/>

02. 1-CLICK LINKEDIN CONNECTION NOTE (< 300 CHARS)

Hi! I built a live interactive Boids + 1.58-bit ternary inference portfolio specifically for Sakana AI's Systems & GPU Infrastructure team in Tokyo: <https://sakana.vaked.dev> (12.8x BitNet compression + Osaurus/Headroom OSS architect). Would love to connect! — P é t e r

03. LINKEDIN COMPANY & LEADERSHIP CONTACTS

- Company LinkedIn: <https://www.linkedin.com/company/sakana-ai/>
- Company Jobs: <https://www.linkedin.com/company/sakana-ai/jobs/>
- Company People: <https://www.linkedin.com/company/sakana-ai/people/>
- Recruiter Search: <https://www.linkedin.com/search/results/people/?keywords=Sakana%20AI%20%28Recruiter%20OR%20Talent%20OR%20Engineer%20OR%20Founder%29>

• David Ha (Co-Founder & CEO, Sakana AI (Former Head of Google Brain Tokyo)):
<https://www.linkedin.com/search/results/people/?keywords=David%20Ha%20Sakana%20AI>
Co-founder leading evolutionary model merging, AI Scientist & collective intelligence in Tokyo.

• Llion Jones (Co-Founder & CTO, Sakana AI (Co-Author of 'Attention Is All You Need')):
<https://www.linkedin.com/search/results/people/?keywords=Llion%20Jones%20Sakana%20AI>
Co-founder & CTO overseeing novel foundation architectures and inference systems.

• Ren Ito (伊藤 隼) (Co-Founder & COO, Sakana AI):
<https://www.linkedin.com/search/results/people/?keywords=Ren%20Ito%20Sakana%20AI>
Co-founder leading operations, strategic partnerships, and Tokyo talent expansion.

04. BILINGUAL MOTIVATION / 志望動機

Sakana AI proves that nature-inspired collective intelligence — evolutionary model merging, The AI Scientist, and composable swarms of efficient specialists — beats brute-force monolithic scaling. My work in 1.58-bit ternary quantization (MLX-QUANT, pureQTern), context compression routing (Headroom PR #1418), and autonomous

agent harnesses (Osaurus, AgentField) is built on the exact same conviction: when you compress foundation weights by $12.80\times$ into $\{-1, 0, +1\}$ and orchestrate them as verifiable swarms, a single workstation or compact GPU cluster becomes an evolutionary laboratory.

Sakana AIが推進する「自然界の群知能に学ぶ進化的モデル統合」と「自律型AIサイエンティスト」は、巨大単一モデルの力任せなスケーリングに対する最も美しい解答です。私が開発してきた1.58ビット三値量子化エンジン（MLX-QUANT / pureQTern、FP16比12.80倍圧縮）やマルチエージェント基盤（Osaurus / AgentField / Headroom PR #1418）は、まさに限られた計算資源から最大の創発的知能を引き出すためのシステム技術であり、東京のSakana AIチームでその推論・分散基盤を極限まで高速化したいと強く願っています。