



web - [borkiss.net](http://borkiss.net)  
 telegram - [@borkiss](https://t.me/borkiss)  
 github - [lubluniky](https://github.com/lubluniky)  
 hugging face - [borkiss](https://huggingface.co/borkiss)  
 x - [@borkiss727](https://x.com/borkiss727)  
 linkedin - [/in/borkiss](https://in/borkiss)

*building the layer between frontier models & production*

Head of AI · Founder @ [fdaa.dev](https://fdaa.dev) · ex-quant

// NUMBERS

**13B**

tokens/quarter through  
agentic pipelines

**8300+**

commits/yr - written by  
agentic systems

**6.5×**

inference speedup  
(WebGPU)

**157★**

on bs-p - flagship open-  
source artifact

// EXPERIENCE

2025 - NOW

**Head of AI - ██████████ (NDA)**

AI strategy, product and infrastructure for the agency: agentic pipelines, content-workflow automation, proprietary AI tooling, vertical-specific safety classifiers.

2025 - NOW

**Founder & Head of Engineering - FDAA ([fdaa.dev](https://fdaa.dev))**

Forward deployed AI acceleration: dev team, open source, client services. Embedding AI engineering right into clients' operations. Workshop business with partners (SRE + Golang).

EX

**AI Experience Advisory - [scope360.io](https://scope360.io) & [cryptomannn.com](https://cryptomannn.com)**

Advisory on AI product experience, agent design and implementation strategy.

PRIOR

**Quantitative Data Analyst / Trader**

Quant analysis, signals, data pipelines, market making. AVX-512 Avellaneda-Stoikov kernel in logit space (bs-p, 157★). The era that built adversarial intuition, fat-tail thinking and a nose for look-ahead bias.



## // SELECTED PROJECTS

---

### Project A – agentic assistant platform (NDA) PRODUCTION

Production assistant on three model tiers; advisor layer engages only on money-impacting actions (~5% of calls). Pentest: every prompt injection repelled.

### Consilium / Heavy Mode MULTI-AGENT

16 agents debate, a composing judge assembles the answer. ~78% ensemble gain from the loop itself; 0.97× wall-clock for 3 concurrent turns.

### WebGPU browser inference IN-BROWSER

LLMs and open-vocab video detection entirely in the browser (WGSL). 6.5× speedup, 724 MB down from 2.4 GB, 91–95 tok/s, 22.6 FPS. 3 live demos: [hf.co/borkiss](https://hf.co/borkiss).

### AgenticBox INFRA

VPS hosting for agents + git-based P2P file sync: ed25519/SPAKE2, Headscale, Rust core, Electron UI. AGPL-3.

### recursive-pi / Ankore RESEARCH

Self-rewriting agents: immortal supervisor + mortal core. A controlled experiment – Goodhart's law reproduced in the lab (1 line of TS in 36 commits).

### UACL PROTOCOL

A communication protocol for heterogeneous agents: hook-push, mailbox channels, unix-socket bus.

### Dynamite / share-plan / xparse TOOLING

Streaming artifact rendering (700–1000 tok/s), plan sharing on CF Workers, X parsing without the official API price tag.

---

## // OPEN SOURCE & ORGS

bs-p – AVX-512 Polymarket MM kernel, C · 157★/35 forks  
rocket-rs – RL in Rust (PPO, A2C)  
primitives – md-to-code · 26 public repos total  
HolyPoly (holypoly.ca) – co-founder, CTO & CAO · Centurion – side

---

## // CHANNELS

borkiss notes (t.me/borkiss\_notes) – agent infrastructure, token economics  
reseaxch – a discord for AI researchers

---

## // STACK

systems Rust · C · Go · Elixir/BEAM  
web TypeScript · Bun · React · Hono  
infra Cloudflare Workers · Terraform · Docker · K8s · Kafka · Headscale  
ai/agents Claude Code · Codex · custom harnesses · WGSL/WebGPU inference · multi-provider orchestration