

Retroactive Funding for JAM Knowledge Base Search Engine

General Project Information

Project Category / Type: Software development

Proponent: FluffyLabs.dev

USDC address: 15fGrDWmFoaApnZXEw1Zg45zuCqpkbMS9YPzczeU9dHseUUu

Requested allocation: 25.000 USDC

- Income taxable: 19% tax on "virtual currencies" tax in Poland:
<https://www.podatki.gov.pl/en/your-e-pit/pit-38-for-2022/>

Polkasassembly discussion: <https://polkadot.polkasassembly.io/post/3296>

Discussion date: 13th July 2025

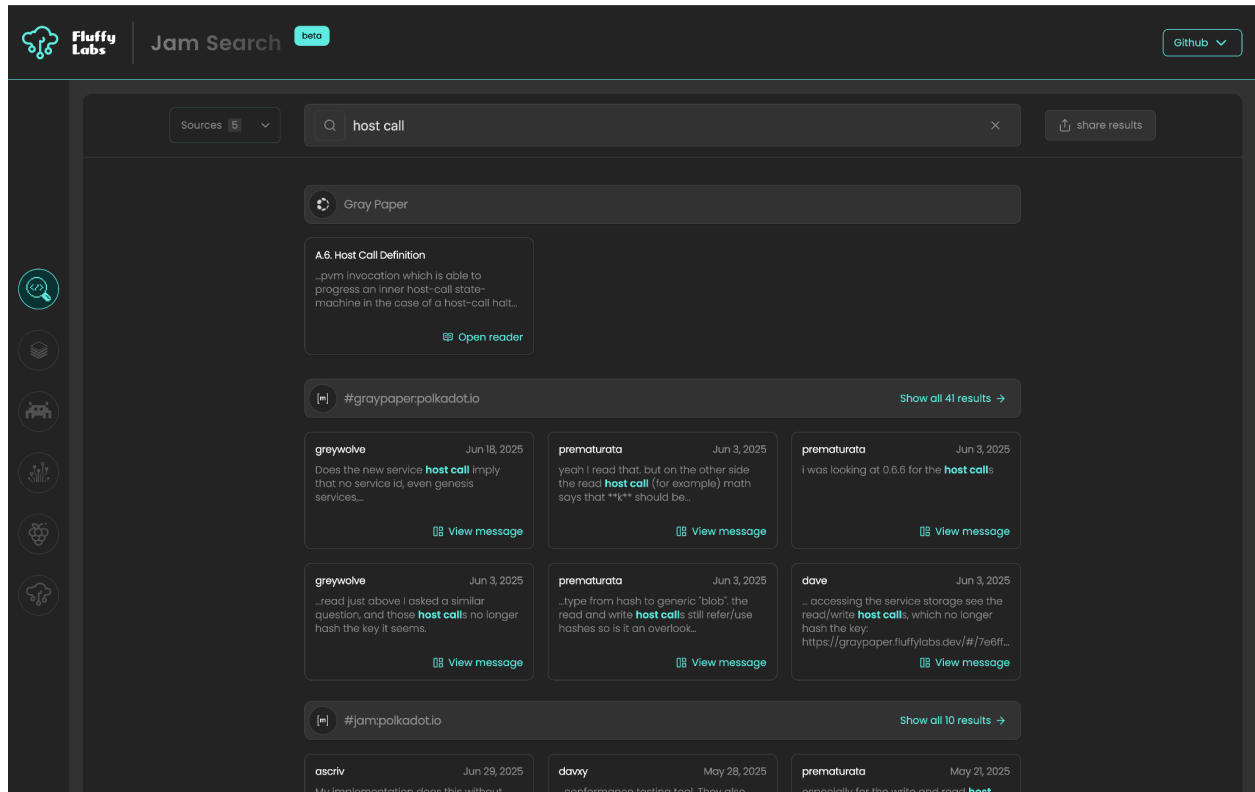
Onchain publish date: scheduled for 20th July 2025

Previous treasury proposals: none

Context

The JAM ecosystem faces a critical information fragmentation problem. JAM-related content is scattered across multiple platforms: Gray Paper specifications, Matrix chat discussions, documentation websites, and GitHub repositories. Over 20 teams participating in the JAM Prize Contest need efficient access to this distributed knowledge base.

We built the JAM Knowledge Base Search Engine to solve the major pain points we identified while working on JAM projects and participating in community discussions.



Problem

While developing JAM-related projects, we identified the following issues:

- **Information fragmentation:** Critical content scattered across Gray Paper, Matrix chats, docs, and GitHub
- **Search limitations:** No unified search across platforms; existing tools don't understand JAM context
- **Context discovery:** Difficulty finding discussions that led to implementation decisions
- **Onboarding barriers:** New developers struggle to discover relevant resources efficiently
- **Knowledge continuity:** Important historical context gets lost in chat logs and archived discussions

Proposal

We propose retroactive funding for the development of the JAM Knowledge Base Search Engine hosted at <https://search.fluffylabs.dev> with open source code available on GitHub (<https://github.com/fluffylabs/jam-search>).

Unlike the JAM prize, which focuses on client implementations, no existing funding covers the creation of ecosystem tools like comprehensive search engines.

Discussion about this kind of solution started on #jam:polkadot.io:



gav

Yes, provided comms are open and easily discoverable by other teams/newcomers, and of course that it doesn't impact the nature of the clean-room rules, then it's fine.

Generally speaking, if discourse happens openly and in easily discoverable channels and it's solely about getting the interop working then there shouldn't be a problem.



Also, I'm cognisant that forcing everything to happen in these two channels isn't ideal. So I've a bit of a plan here:

1. Get some sort of automation to periodically archive these two channels.
2. Host the archive in one or more obvious places (graypaper.com, jamcha.in, others?)
3. Request that others who want to have discourse away from these channels offer up the same thing and provide linking/hosting on these central spots for it.
4. Have some AI guy train an "official" AI model periodically on the content of these channels as well as any lectures of mine, the Gray Paper &c.
5. Provide the model as a download for anyone to use with e.g. Jan.

As long as *all* non-official channels data gets into this central easily-searchable/discoverable knowledge base in a timely fashion, then I don't see a huge problem with expanding out from the two official channels.



Of course I won't necessarily be monitoring other channels.

+5

Available Solutions

Existing search solutions have significant limitations:

- **GitHub Search:** Limited to code repositories, doesn't index community discussions
- **Matrix Search:** Restricted to individual rooms, no cross-room or semantic capabilities
- **Documentation Site Search:** Each site has separate, limited search functionality
- **General Search Engines:** Cannot index private Matrix rooms or understand JAM-specific context

The JAM Knowledge Base Search Engine is the first solution providing unified, intelligent search across all JAM content sources.

Features

The JAM Knowledge Base Search Engine features:

- **Multi-source search:** Unified search across Gray Paper, Matrix chats, Discord channels, JamCha.in documentation, and GitHub
- **Advanced search modes:** Strict text matching, fuzzy search, and AI-powered semantic search
- **Real-time synchronization:** Automatic updates from all content sources
- **Advanced chats filtering:** Filter by source, date ranges, channels, and GP versions
- **RESTful API:** Integration endpoints for other ecosystem tools
- **Responsive interface:** React app optimized for desktop and mobile

Deliverables

Web Application: <https://search.fluffylabs.dev>

Open Source Code: <https://github.com/fluffylabs/jam-search>

Backend API: RESTful endpoints with comprehensive documentation

Data Pipeline: Automated collection systems for all JAM content sources

Adoption

The search engine is already being used by JAM development teams and has proven effective in reducing time spent searching for technical information across multiple platforms.

Timeline

This proposal seeks retroactive funding for the development of the JAM Knowledge Base Search Engine, which was developed during March-May 2025 based on git commit history:

- **March 2025:** Initial project setup and foundation
- **April 2025:** Core development - API, database, frontend, Matrix integration
- **May 2025:** Advanced features - semantic search, GitHub integration, optimization
- **June 2025:** Discord integration

Budget

The JAM Knowledge Base Search Engine development is complete, and we are seeking retroactive funding for 250 hours of work, covering full-stack development, AI integration, and deployment.

Budget Breakdown

Task	Time (Hours)	Rate (\$/hour)	Cost
Backend Development	API development, database design, Matrix integration, Discord integration, Websites integration, Graypaper indexing	120 hours	\$100
Frontend Development	Design, React app, UI components, search interface	80 hours	\$100
AI Integration	Semantic search with OpenAI embeddings	30 hours	\$100
DevOps & Deployment	Docker setup, database optimization, production deployment	15 hours	\$100
Project Management	Planning, QA, documentation, coordination	5 hours	\$100

Total Budget:

Developer Hours: 245 hours × \$100/hour = \$24,500

Project Management Hours: 5 hours × \$100/hour = \$500

Total: \$25,000

This budget reflects the full scope of work required to develop and deliver the open-source JAM Knowledge Base Search Engine.

Infrastructure cost

- Server - 7 USD per month
- Database - 19 USD per month

Total - 26 USD per month. The infrastructure cost is covered until July 2026

Proponent

USDT Address: 15fGrDWmFoaApnZXEw1Zg45zuCqpkbMS9YPzcieU9dHseUUu

Requested Allocation: 25.000 USDC

Governance Referenda Origin Call: Small spender

Contact Information

Team Member: Krystian Fras

Email: krystianfras95@gmail.com

Matrix: @krystian50:matrix.org

GitHub: <https://github.com/krystian50>

Team

The **JAM Search** was built by the following developers as part of the **FluffyLabs** collective, a group involved in the development of decentralized tools for the Polkadot ecosystem:

- **Krystian Fras**

Krystian led the complete development of the application, from initial concept through production deployment. He's a PBA Alumni ([certificate](#)) and developer of the PVM Debugger ([treasury proposal](#)).

- **Piotr Wojciechowski**

Piotr was responsible for search engine implementation, data mapping and AI capabilities. He took a part in both frontend and backend development.

- **Marcin Raczyński**

Marcin designed the application and was responsible from maintaining a proper UX through the whole implementation process.

- **Tomasz Drwięga**

Tomasz was a core developer at Parity Technologies, where he worked on various core parts of Substrate. Now, as founder of Fluffy Labs contributing to Polkadot Ecosystem, including leadership of the JAM implementation. Tomasz brought the initial search application idea and guided through the process to achieve a valuable application.