Proto-OKN: Cross-Knowledge Graph Queries

Natural Language Queries using MCP Servers

Peter Rose (pwrose.ucsd@gmail.com)

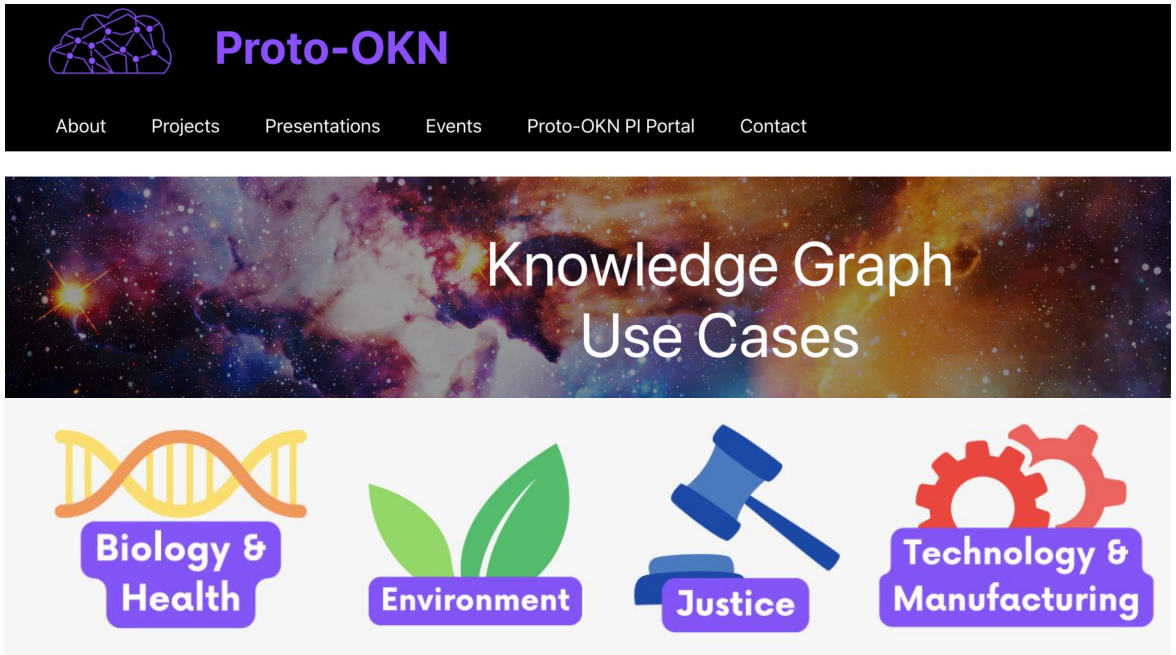
SPOKE TEAM
PI: Sergio Baranzini (UC San Francisco)
Peter Rose (UCSD), Sui Huang (ISB), Scooter Morris (UCSF),
Angela Rizk-Jackson (UCSF), Yongmei Shi (UCSF), Sam Gebre (NASA),
Amanda Saravia-Butler (NASA), Kirill Grigorev (NASA),
Aenor Sawyer (UCSF/NASA), Charlotte Nelson (Mate Bioservices)

WOBD TEAM (The Scripps Research Institute, TSRI)
Andrew Su, Benjamin Good

FRINK TEAM (Renaissance Computing Institute, RENCI)
Christopher Bizon, Jim Balhoff, Yaphet Kebede

NSF Awards 2333819, 2535091

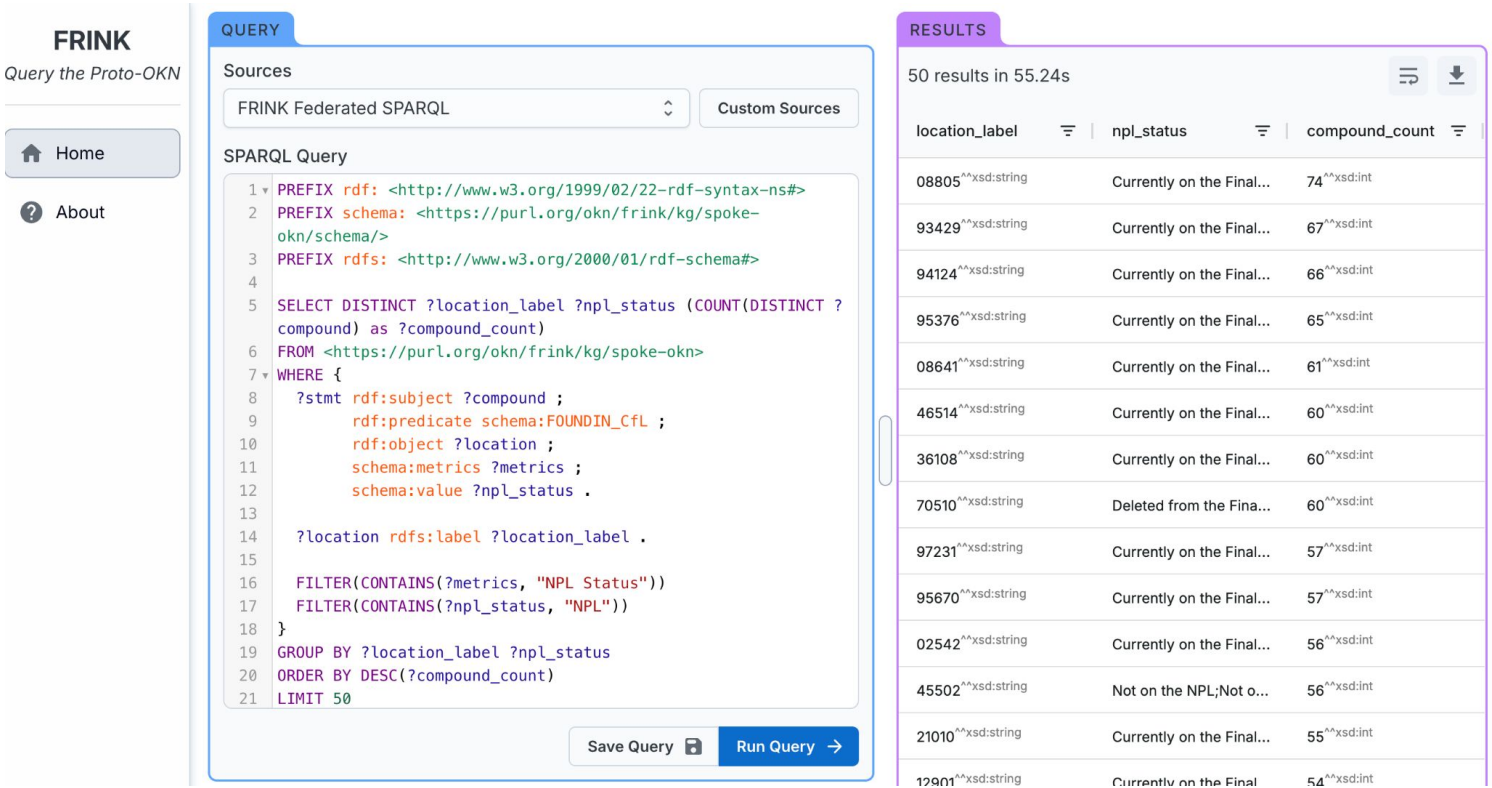
NSF Proto-OKN Program



“Open knowledge network (OKN) is a publicly accessible, interconnected set of data repositories and associated knowledge graphs that will enable data-driven, artificial intelligence-based solutions”

https://www.proto-okn.net/projects/

FRINK: SPARQL Query of the Proto-OKN KGs

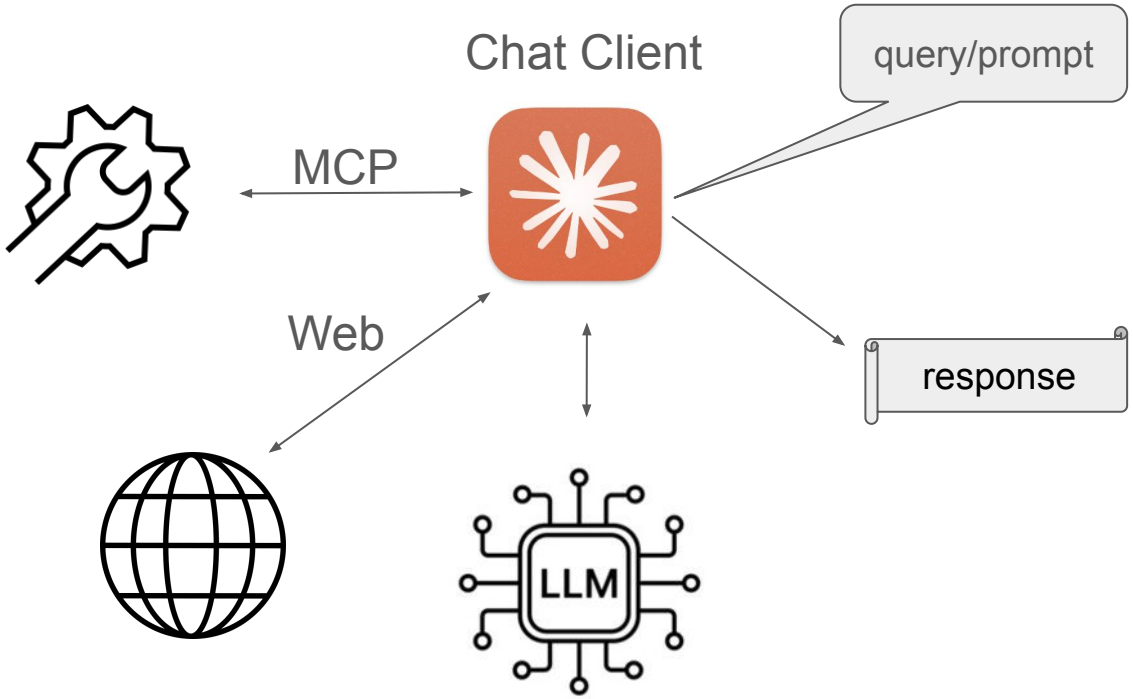


SPARQL queries are verbose and complex

Few users have SPARQL experience

https://frink.apps.renci.org/

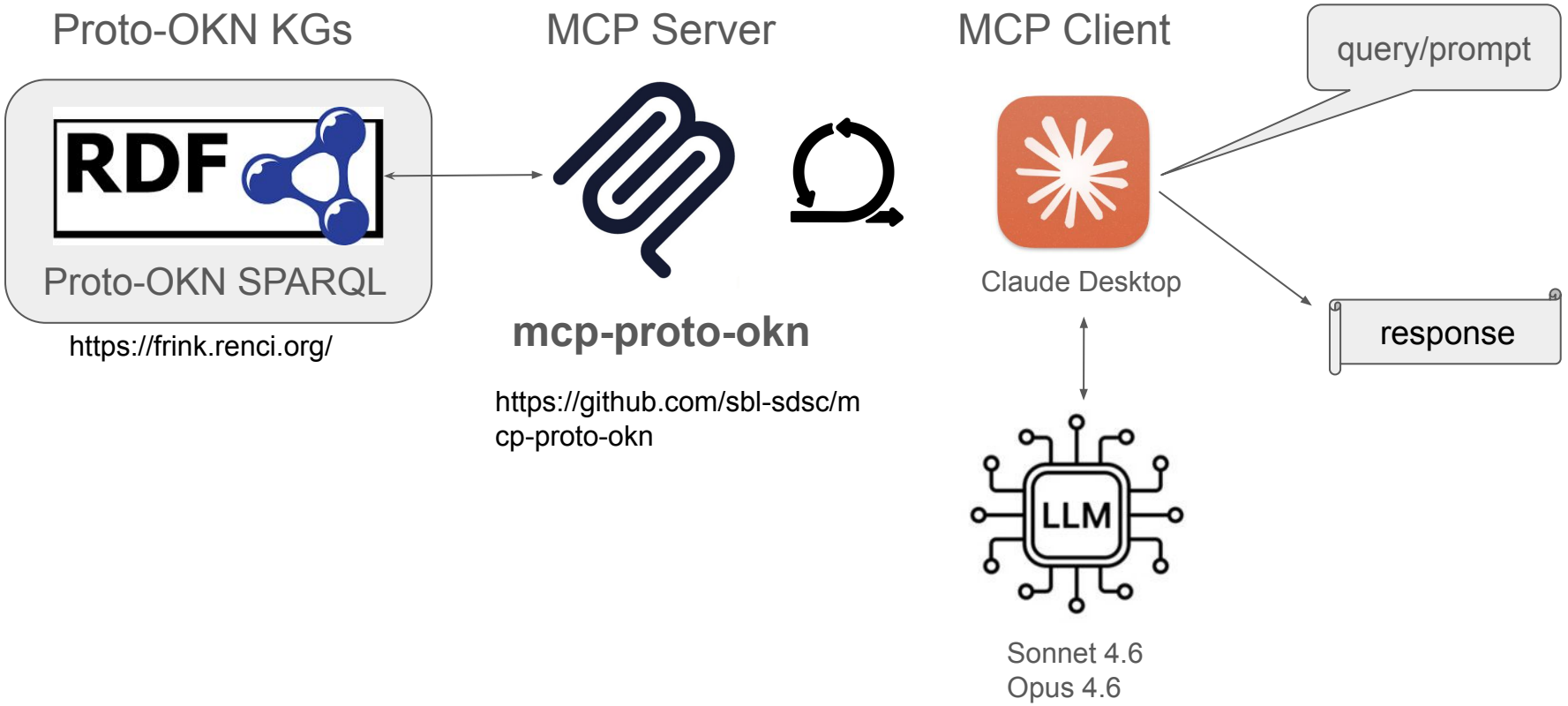
AI Assistants and Tools



MCP (Model Context Protocol) is an open protocol that standardizes how applications provide context to large language models.

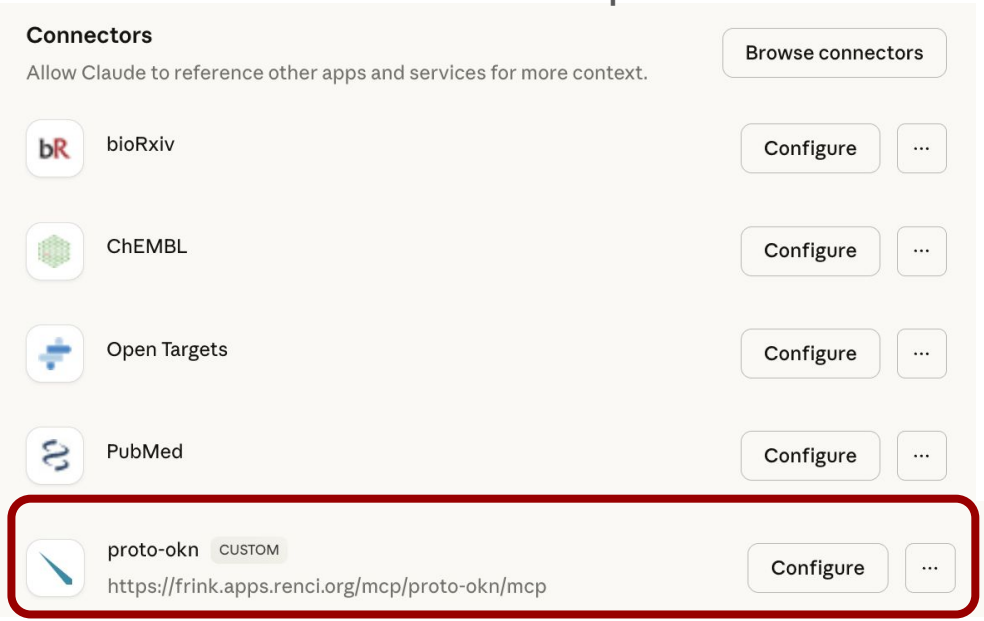
https://modelcontextprotocol.io/

MCP-Proto-OKN - Natural Language Queries

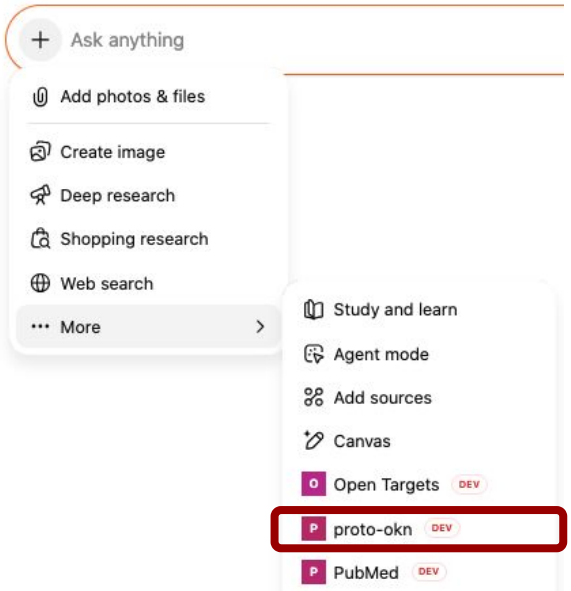


Setup mcp-proto-okn in AI Assistant

Claude Desktop



ChatGPT



Configuration instructions:
https://github.com/sbl-sdsc/mcp-proto-okn?tab=readme-ov-file#installation-and-configuration

Demos

1. Proto-OKN KGs Descriptions

Generate a table of all Proto-OKN Knowledge Graphs with two columns: “KG Name” and “Description”

2. Single KG query

@spoke-genelab: Give a high-level overview of this knowledge graph, including its main entities, relationships, and purpose

3. Perfluorooctanoic acid (PFOA) safety profile - cross-KG

https://github.com/sbl-sdsc/mcp-proto-okn/blob/main/docs/examples/pfoa-safety-profile.md

4. Spaceflight disease risks - cross-KG + cross-MCP

https://github.com/sbl-sdsc/mcp-proto-okn/blob/main/docs/examples/space-flight-disease-relationships.md

5. Spaceflight microbiome investigation - cross-KG + cross-MCP

https://github.com/sbl-sdsc/mcp-proto-okn/blob/main/docs/examples/spaceflight_microbiome_cross-graph_investigation.md

6. Full list of examples

https://github.com/sbl-sdsc/mcp-proto-okn?tab=readme-ov-file#example-queries

Tips for Writing Better Prompts

- Be specific, provide context, and clearly state your goal
- Break complex tasks into steps
- Constrain the scope (e.g., specific KG) and anchor to known entities
- Ask for the schema to explain the KG's structure
- Request evidence, assumptions, or intermediate reasoning
- Redirect it when it goes off track and iterate
- Give clear feedback on plots or diagrams
- Ask it to improve your prompt
- Ensure web search is off
- Specify the desired output format (table, csv, plot, etc.)
- Use the prompt "Create chat transcript" to create a document with prompts, responses, and model version that you can save as a .md or .pdf file (see examples)