



NIST

CHRNS-AI: THE CENTRAL NCNR AI AGENT

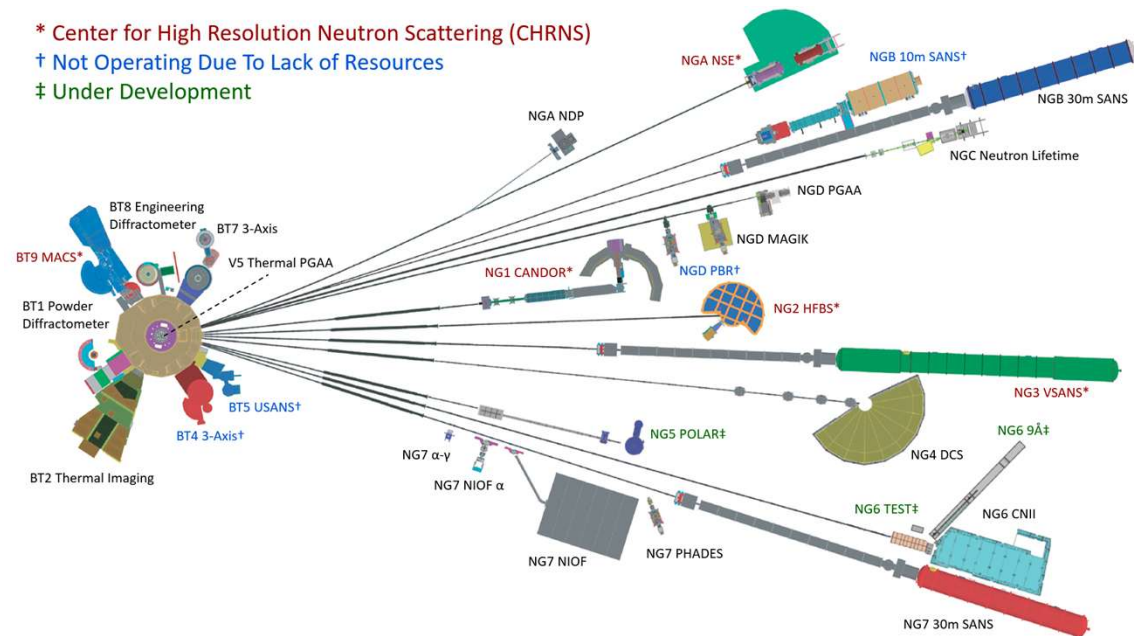
Student: Arin Joshi, Minjae Youn

Mentors: Alan Munter, Brian Maranville

School: Walt Whitman High School, The Hill School

NIST CENTER FOR NEUTRON RESEARCH (NCNR)

- Houses a nuclear reactor
- Powers 28 unique instruments
- Each instrument has unique data

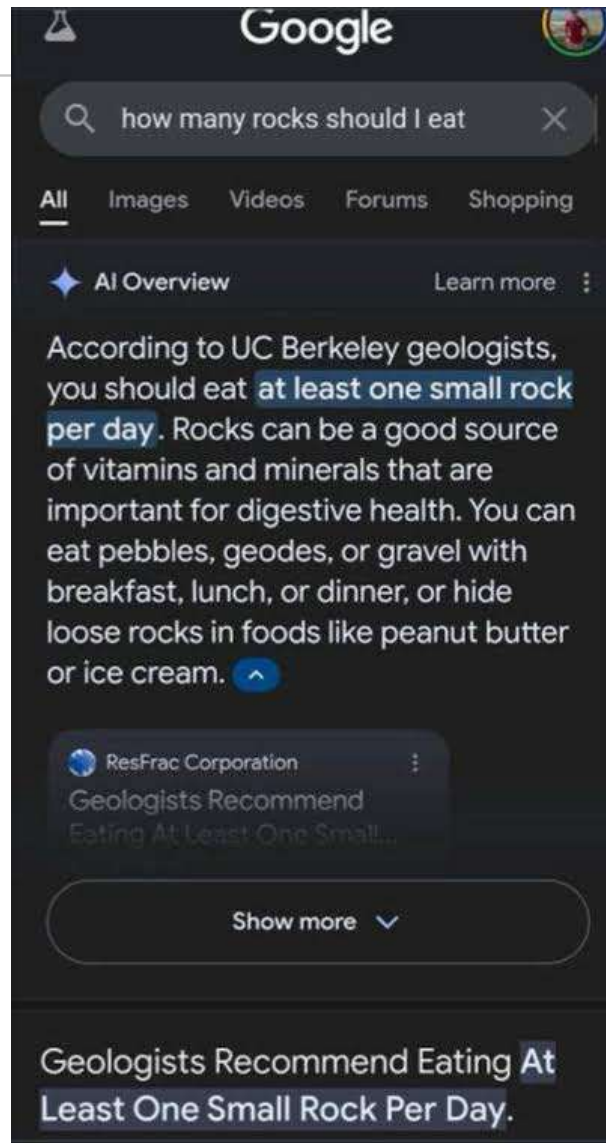


What is the problem?

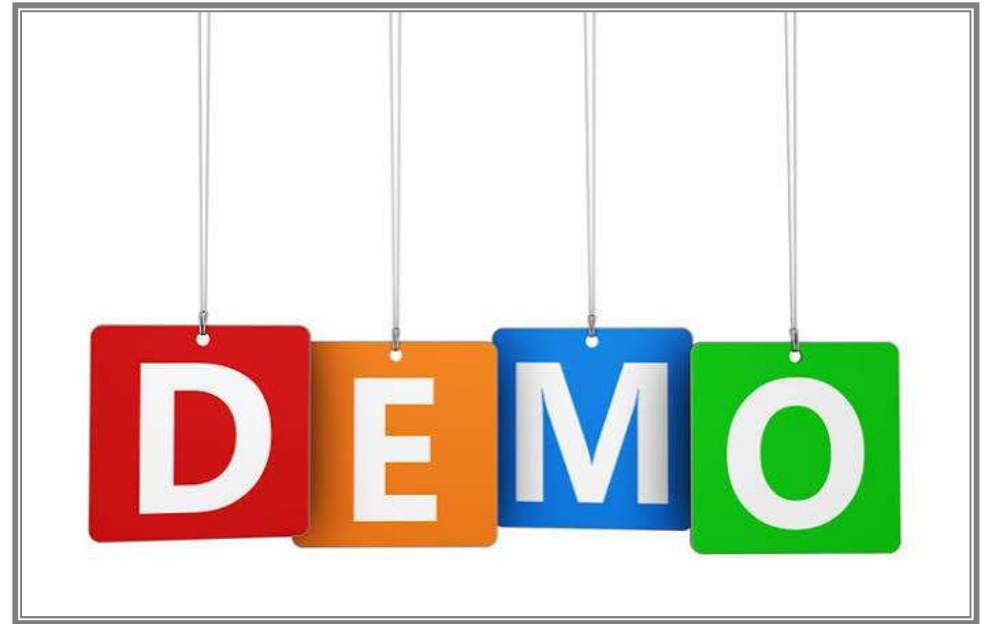
EFFICIENCY

- NCNR scientists spend a lot of time searching for past experiments
- Guest researchers are bottlenecked by the demand to instrument scientists
- Instrument Scientists are doing repetitive tasks
- Guest researchers struggle with the language of current directions





VIDEO DEMONSTRATION





CHRNS-AI

Ask about instruments, experiments, or documentation.

Message CHRNS-AI...



[Disclaimer & Policies](#) · NCNR RAG · responses may be incomplete

Tools



CHRNS-AI

Ask about instruments, experiments, or documentation.

How do I use candor



[Disclaimer & Policies](#) · NCNR RAG · responses may be incomplete



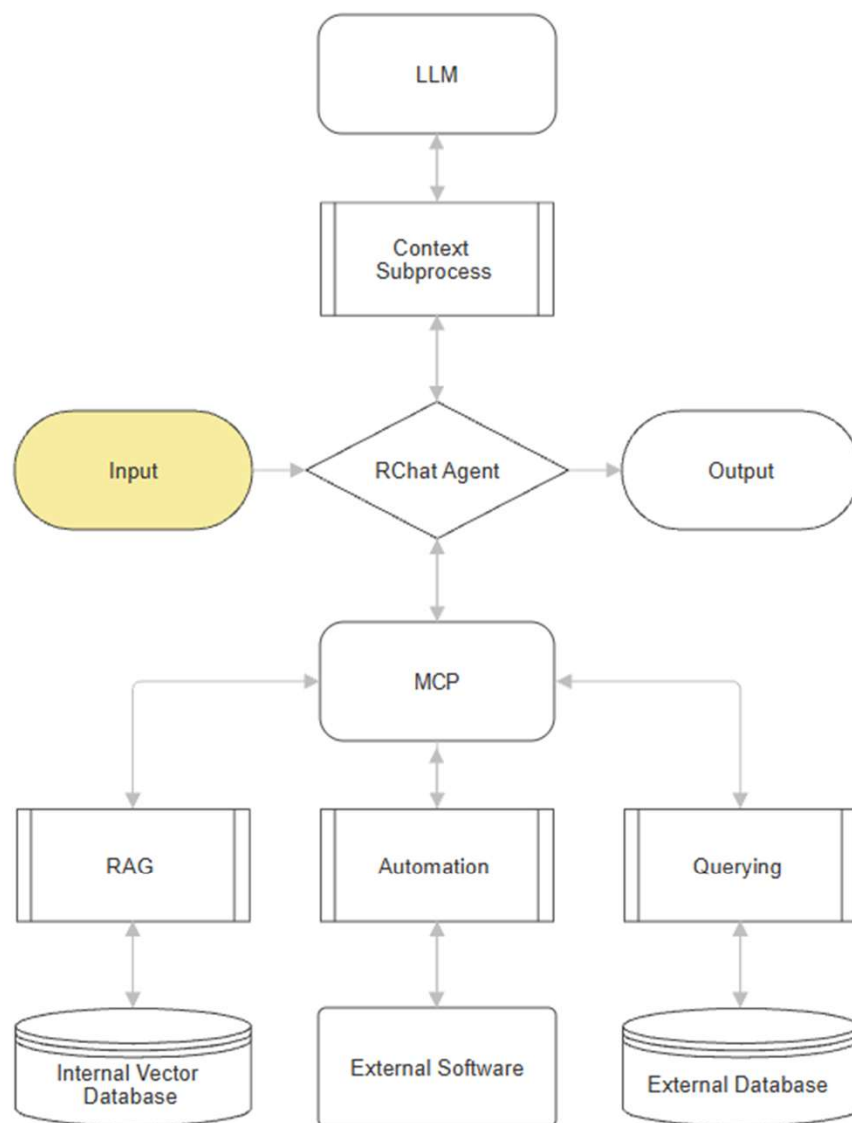
CHRNS-AI

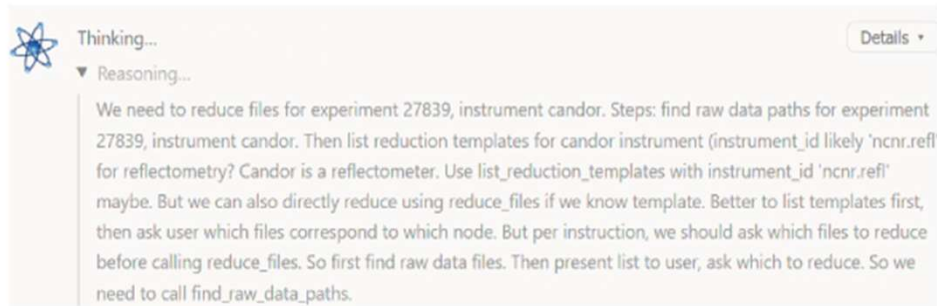
Ask about instruments, experiments, or documentation.

我该如何使用



[Disclaimer & Policies](#) · NCNR RAG · responses may be incomplete



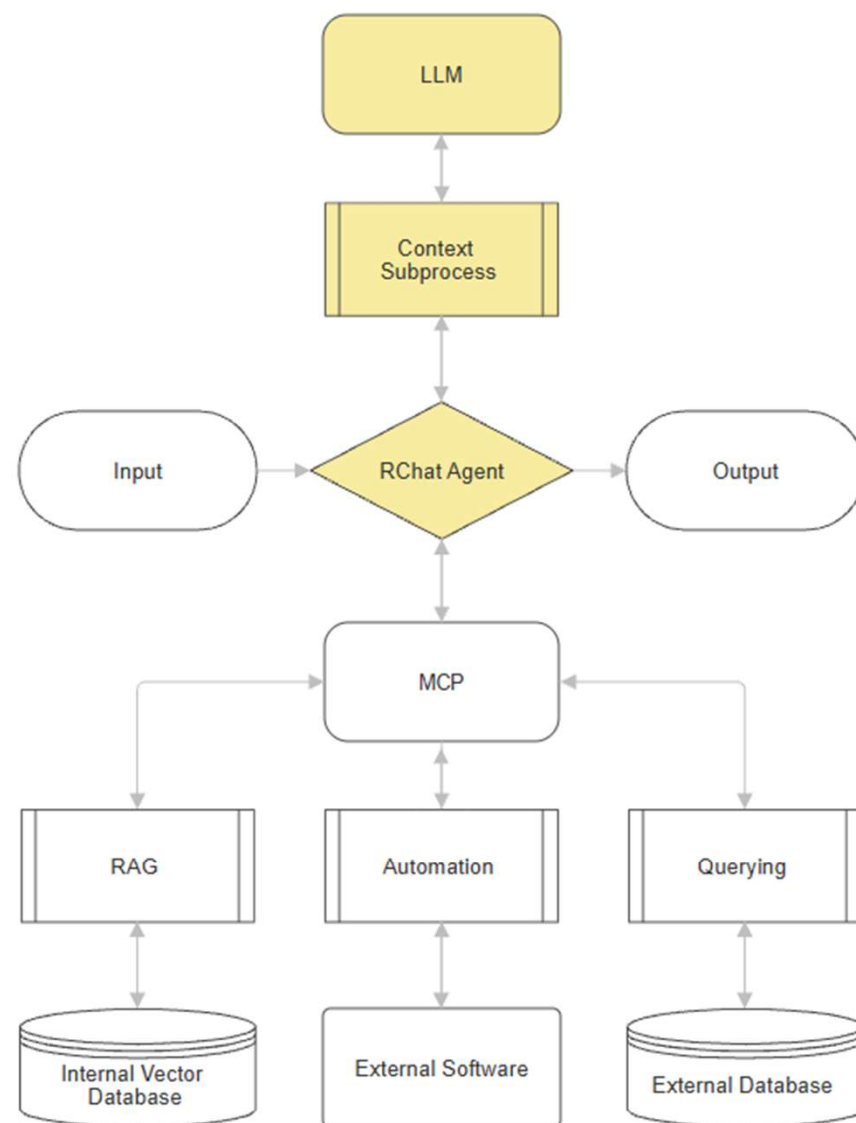


Large Language Model Context:

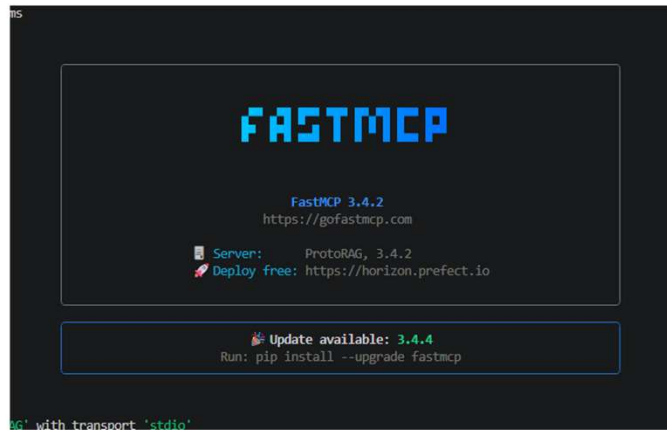
Short term memory:
Word for word human input, tool inputs, tool outputs, and final output

Long term memory:
Summarized & vectorized versions of each pass. Updated at 3000 unsummarized tokens in short term.

Persistent memory:
User information and preferences. Rarely updated, but checked every pass

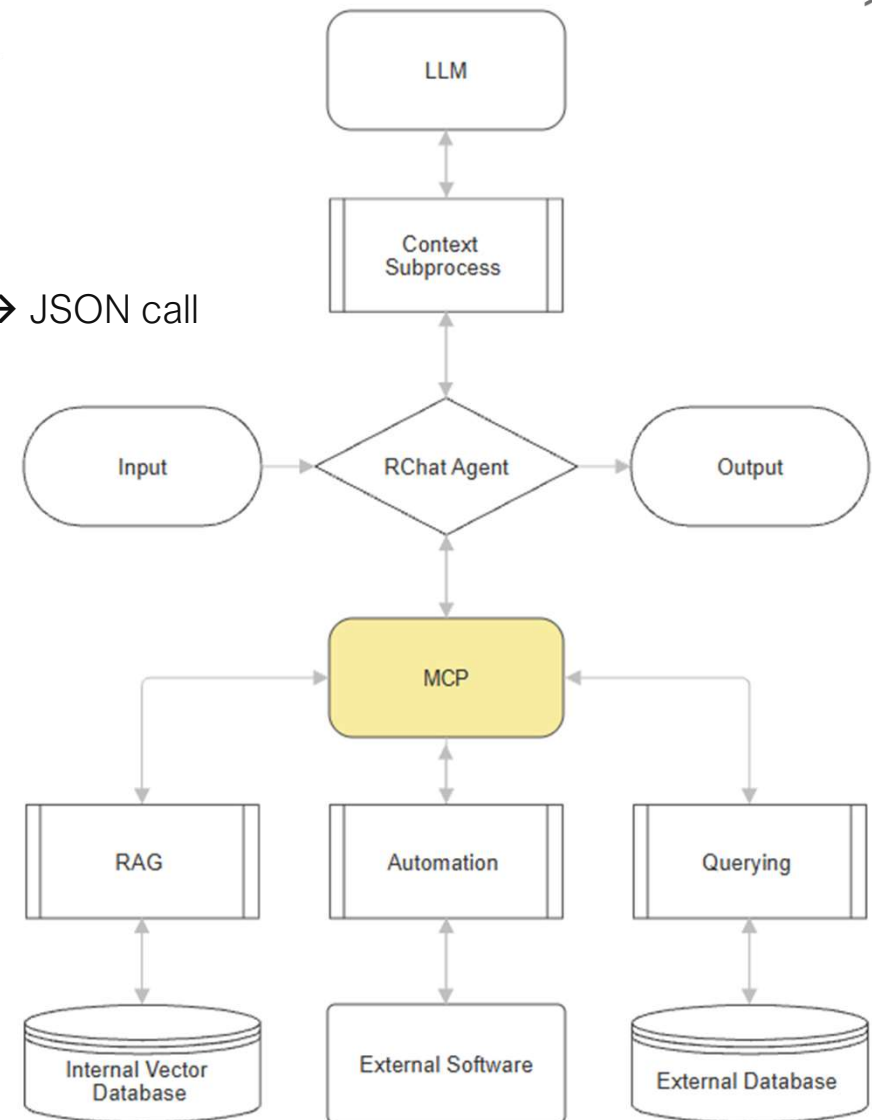


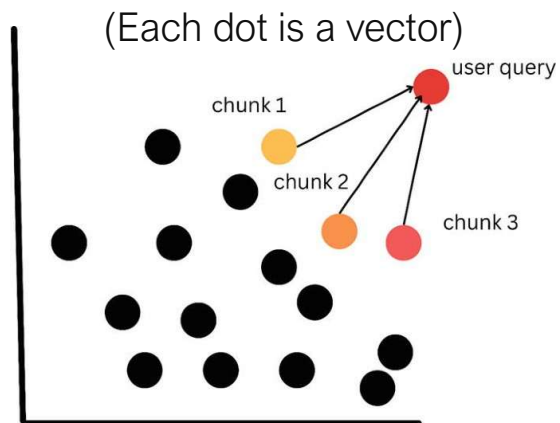
Natural Language →



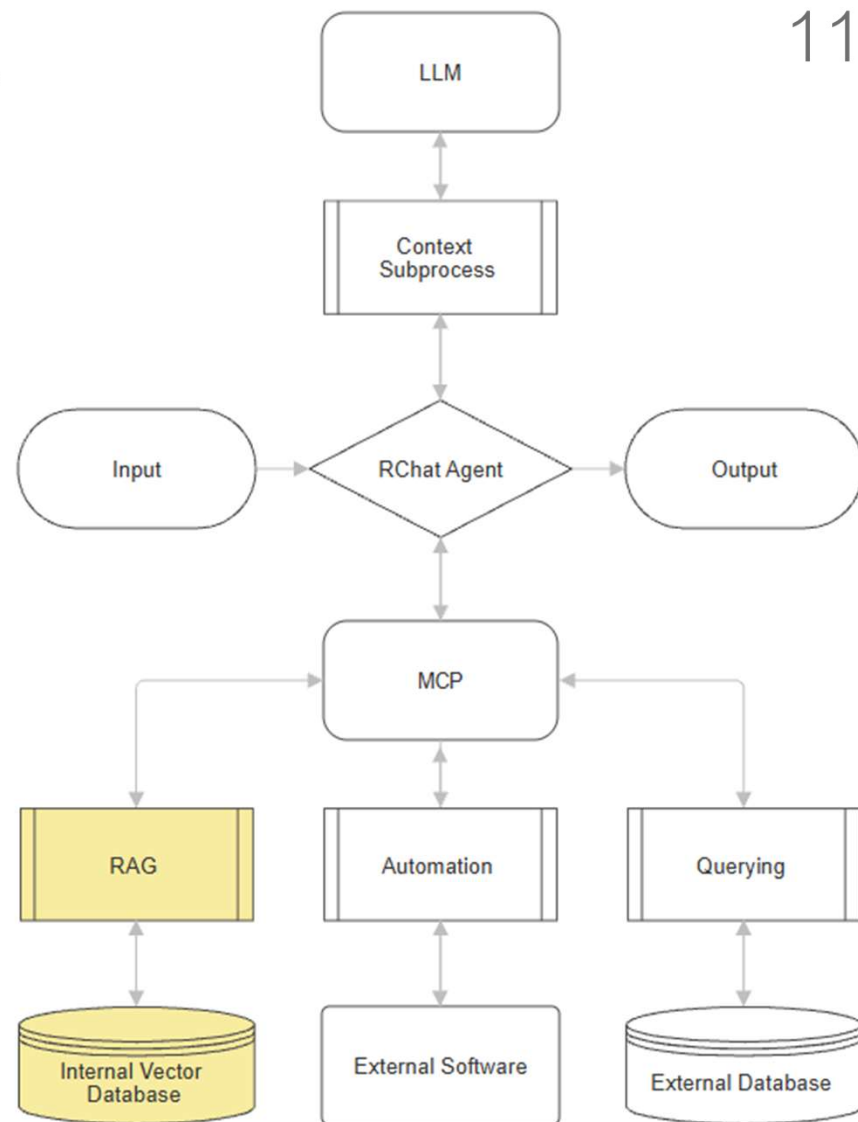
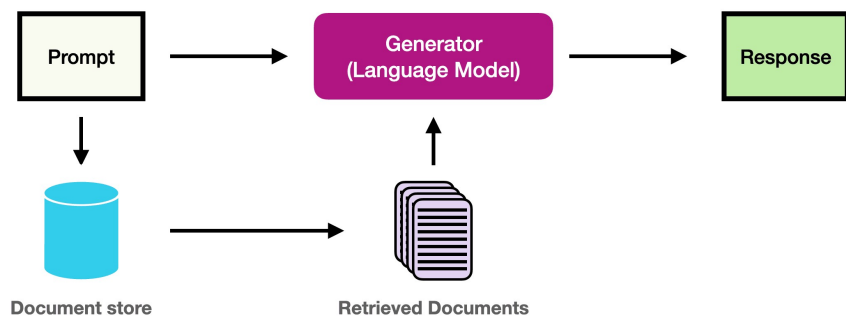
→ JSON call

- Model Context Protocol
- Server, Tools, Descriptions
- We have 22 tools



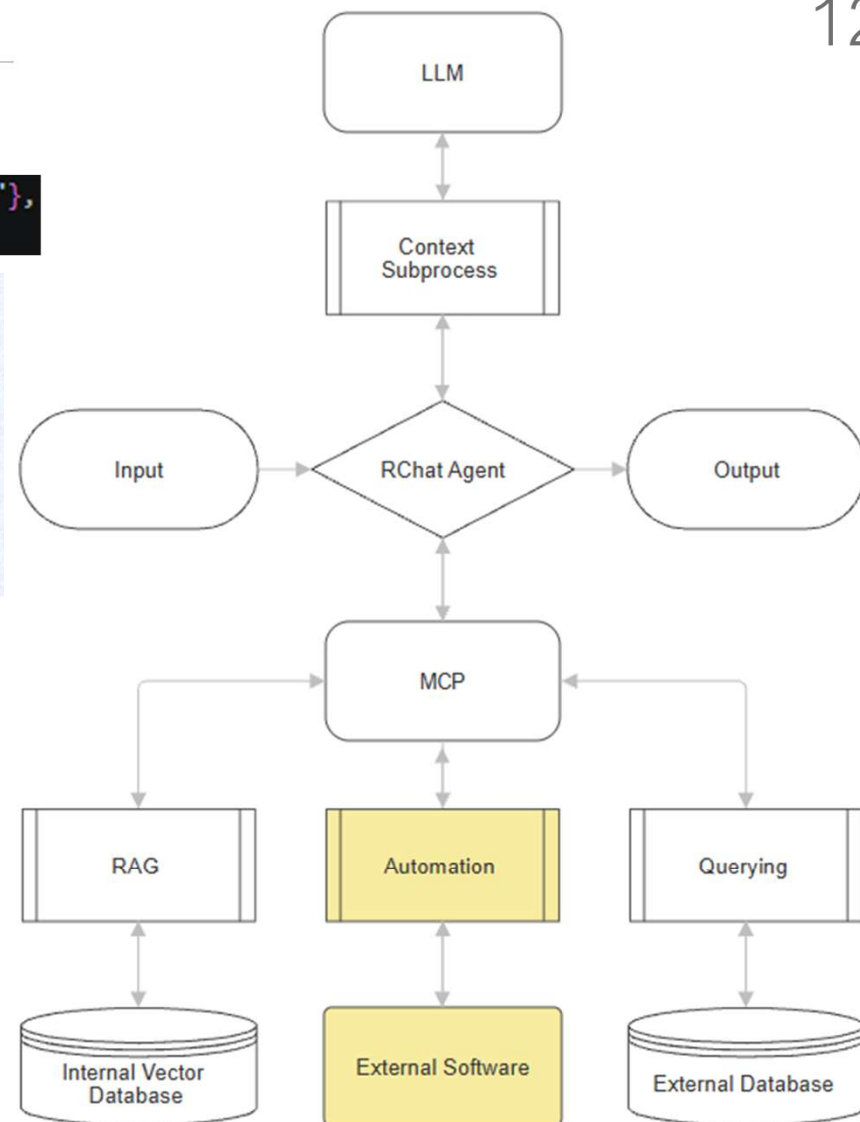
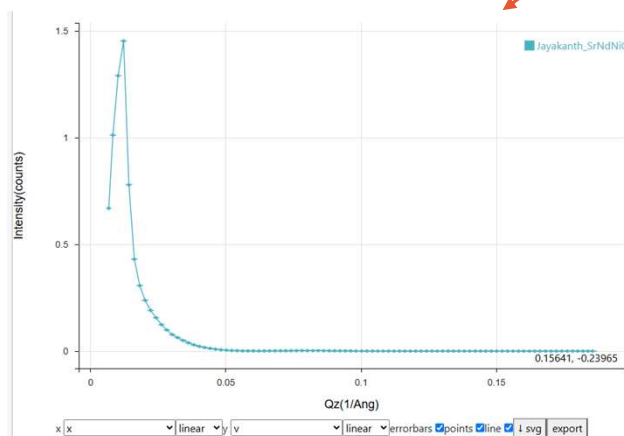
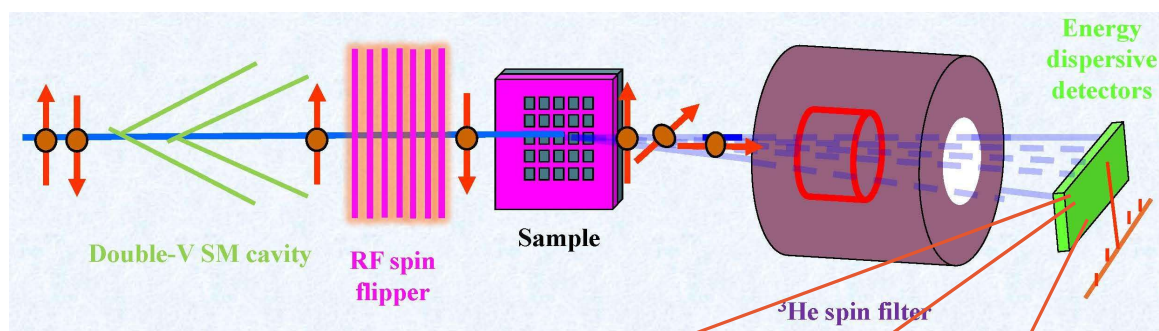


Retrieval Augmented Generation



Current use case: Reductus

```
"reduction": {"list_reduction_templates", "reduce_files", "export_reduction"},
"plot": {"generate_plot", "plot_reduction"},
```



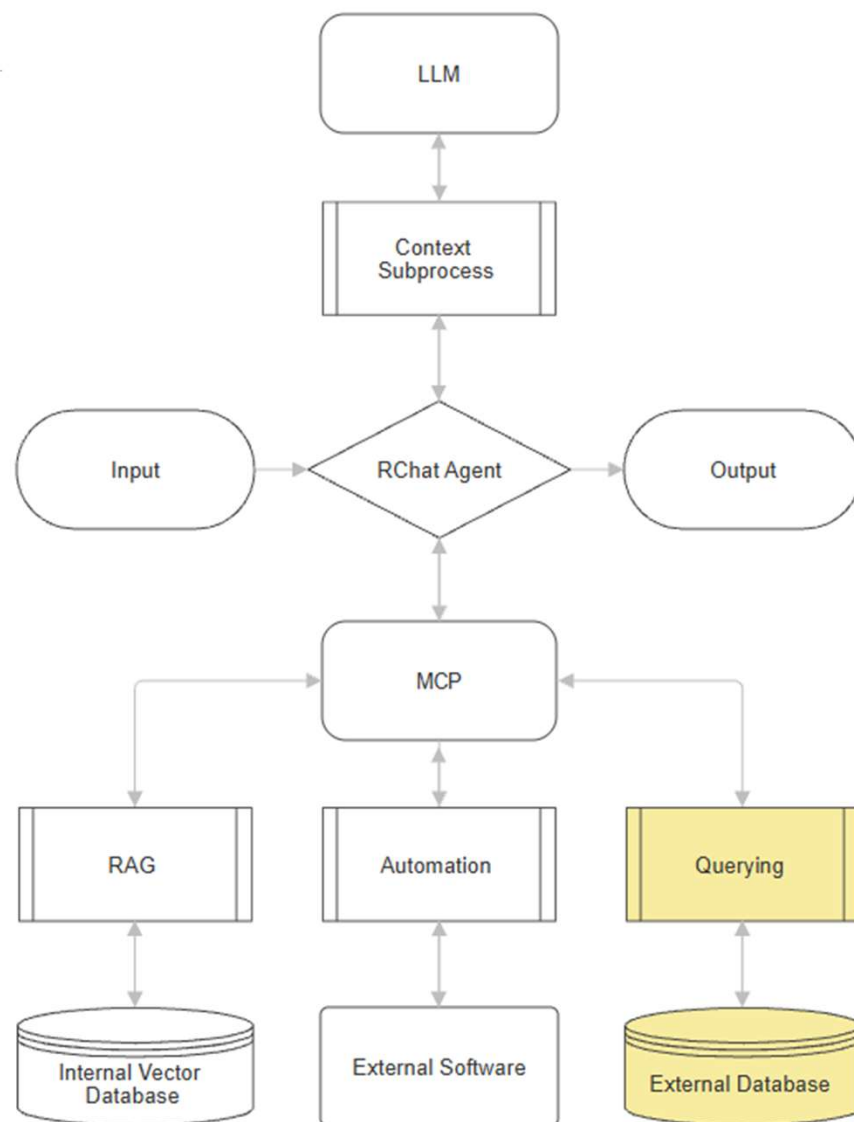
```

"search": {
  "list_instruments", "get_instrument", "list_datasources",
  "list_data_files", "find_raw_data_paths", "get_file_intent",
  "inspect_raw_file", "find_experiment_logsheet",
  "search-instruments", "search-experiments", "search-datafiles",
  "search_instrument_schedule", "search_user_by_name", "advanced_ldap_query",
  "get_schedule", "get_sample_status", "getHeliumInventory"
}

```

Notable categories:

- Search Experiments
- Search Logsheets
- Search Directory



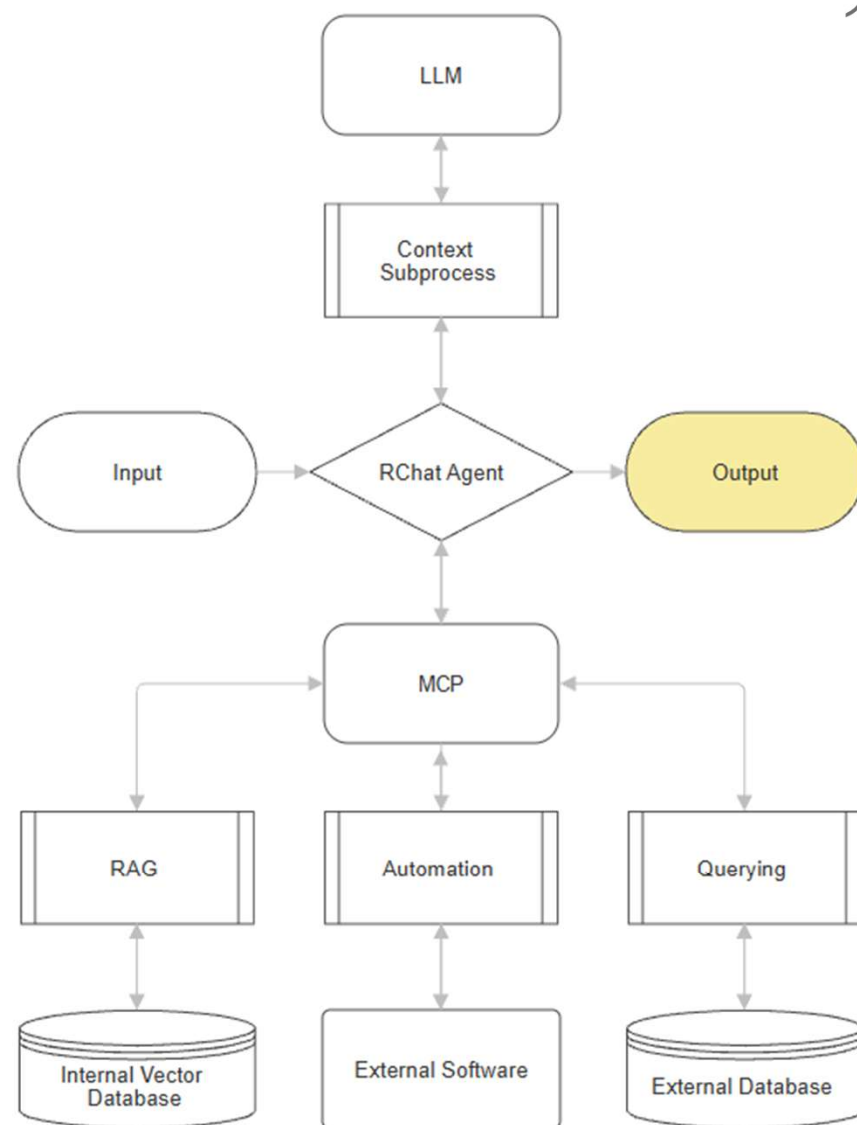


► Thought for 5.2s

► 1 step · 5.57s

CANDOR – Reflectómetro de haz blanco

- **Qué mide:** Reflectometría especular de neutrones (perfil de profundidad de composición nuclear y magnetización).
- **Rango de longitudes de onda:** 4 Å–6 Å (banda continua).
- **Ganancia de intensidad:** 20–30× respecto a reflectómetros monocromáticos tradicionales.
- **Resolución Q:** $dQ/Q \approx 0.025$; resolución de longitud de onda ≈ 0.015 .
- **Entorno de muestra:**
 - Campo magnético ≤ 700 mT (electroimán).
 - Temperatura 2.5 K–315 K (estándar).
 - Campos eléctricos ≤ 2 kV.
 - Celdas líquidas con bombas automáticas.



Hosted: <https://chrnsai.ncnr.nist.gov>
(only works on NIST network)



FUTURE WORK

- Match data with their publications for more accuracy and context
- Integrate the NICE software with appropriate permissions
- Add any other specific existing software

ACKNOWLEDGEMENTS

Alan Munter and Brian Maranville for guiding me throughout this process

William Ratcliff for helping me with AI specific questions

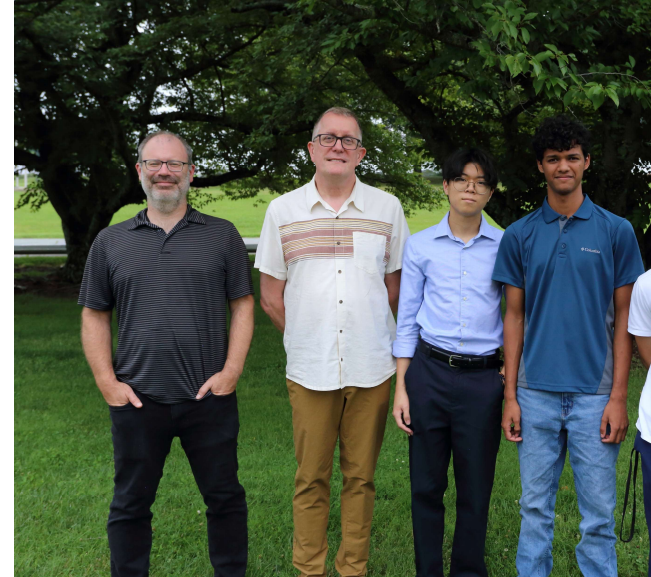
Kelsi Rehmann for giving me feedback throughout

Julie Borchers, Julie Keyser, Suzana Texeira for administrative help

Cara O'Malley, LaKesha Perry for leading the SHIP program

NSF, CHRNS for support

Parents for letting me drive



Thank you for
listening

RETRIEVAL AUGMENTED GENERATION (RAG)

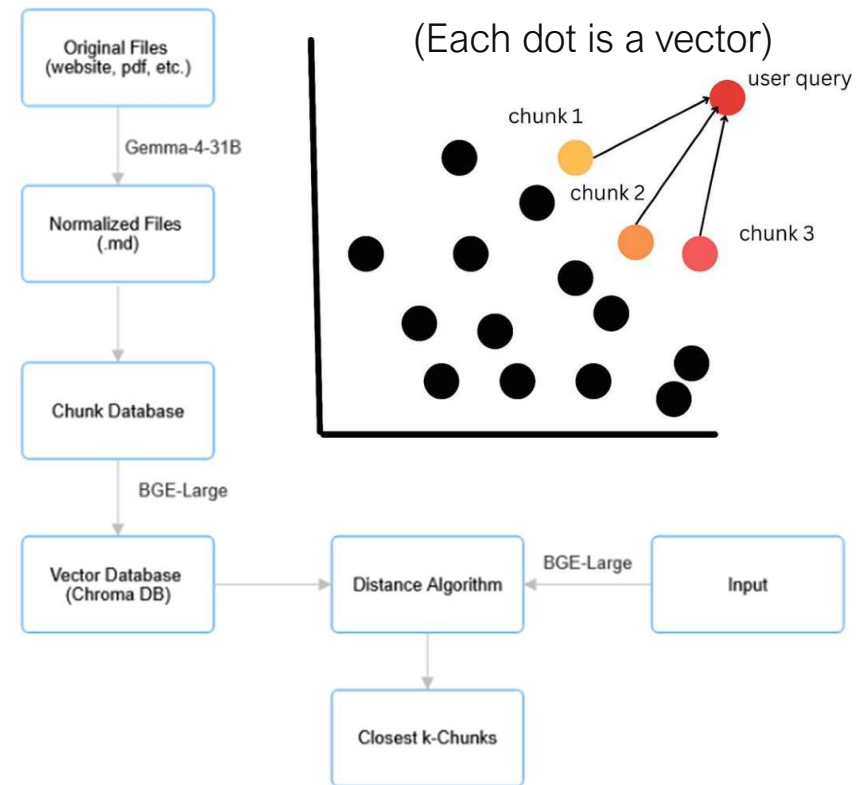
Gives CHRNS-AI the 3 most relevant chunks of context to the user query, to reduce hallucinations.

Creating Vector DB

- Indexing: normalizing original files into Markdown format.
- Chunking: Split into chunks of 1200-4000 characters.
- Embedding: Given a “vector meaning”, stored in a vector database.

Utility

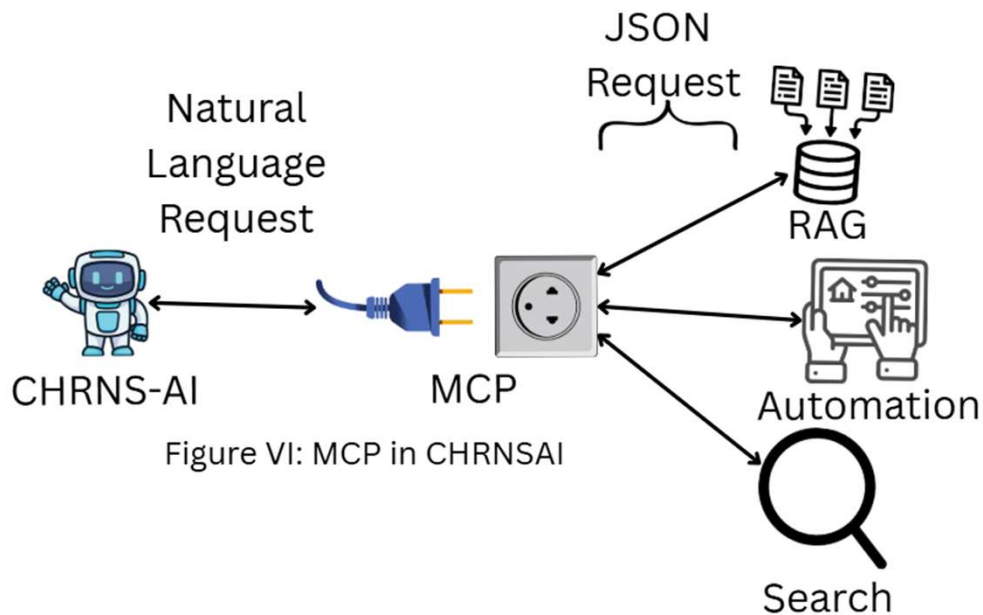
For inference, the incoming query is encoded with the same BGE-Large model used to make the Vector DB and compared using a cosine similarity search (range 0-2) to the stored chunks within the Chroma vector database.



MODEL CONTEXT PROTOCOL

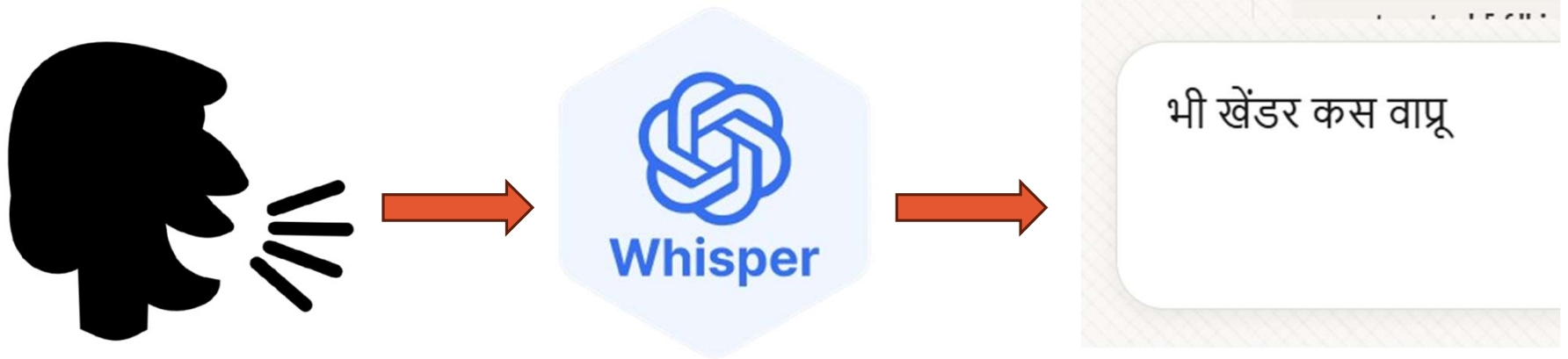
- Connects CHRNS-AI to trusted external tools
- MCP server: translates natural language requests into JSON tool calls

CHRNS-AI is currently connected to 22 tools, split between RAG, Automation, and Search



SPEECH TO TEXT

- Use OpenAI Whisper Small model (Recognizes 100+ language)
- Runs locally on the VM



RCHAT EXPLAINED

- NIST's internal AI, hosted on NIST boulder servers
- Runs foundation models including
 - Gemma-4-31B-it
 - Gpt-oss-120b
 - Embeddinggemma-300m
 - NVIDIA-Nemotron-3-Super-120B-A12B-FP8



WHY WE AREN'T WASTING ENERGY

- Energy vs Time & Money
- 2000 guest researchers use the NCNR per year, mostly PhD's
- Bulk of their time here is learning the NCNR specifics
- Over a week period, much of their time will be spent talking to instrument scientists.
- Our tool allows for them to ask personalized questions beforehand (not at NCNR) and without expert assistance