

DEMOCRATIZING IT NOC DOCUMENTATION/PROCEDURES USING LOCAL AI/LLM

UTILIZING LOCAL AI LLM MODELS AND DOCUMENT CONTEXT
REPOSITORIES FOR GREATER EFFICIENCY



Preface

- High Level presentation
- Not technical deep dive into AI, Gen AI, Machine Learning, Models et al.
- Not a discussion on current end user AlaaS(AI as a Service) tools such as ChatGPT, Copilot, Canva, Grammarly, Decktopus, Docalysis, Dall-E, Midjourney that individually can be used in a similar manner based on individual use case.
- Simple demonstration of using current available AI tools securely and locally to assist in network and help desk staff work in their daily roles.
- Can be applied to any organization's IT or business units.

Converse with your data.

- Utilize your current document repository for additional efficiency.
- Democratize your on-premise static data for daily support use.
- Using local and secure AI tools to query and “converse” with your data in a new manner.
- Large Language Models used on-premise with your documents for local context chat use for easier research and answers.
- Create a chat model for all your documents to be referenced currently in local chat bot environment.

Data democratization

- Allows everyone in an organization to have access to the data they need, and the necessary tools needed to use this data effectively. In short, data democratization enables better business decisions.
- Is the process of making data accessible to everyone in an organization, along with the tools and training needed to understand it. The goal is to enable non-specialists to gather and analyze data without needing outside help.

To democratize data, organizations can:

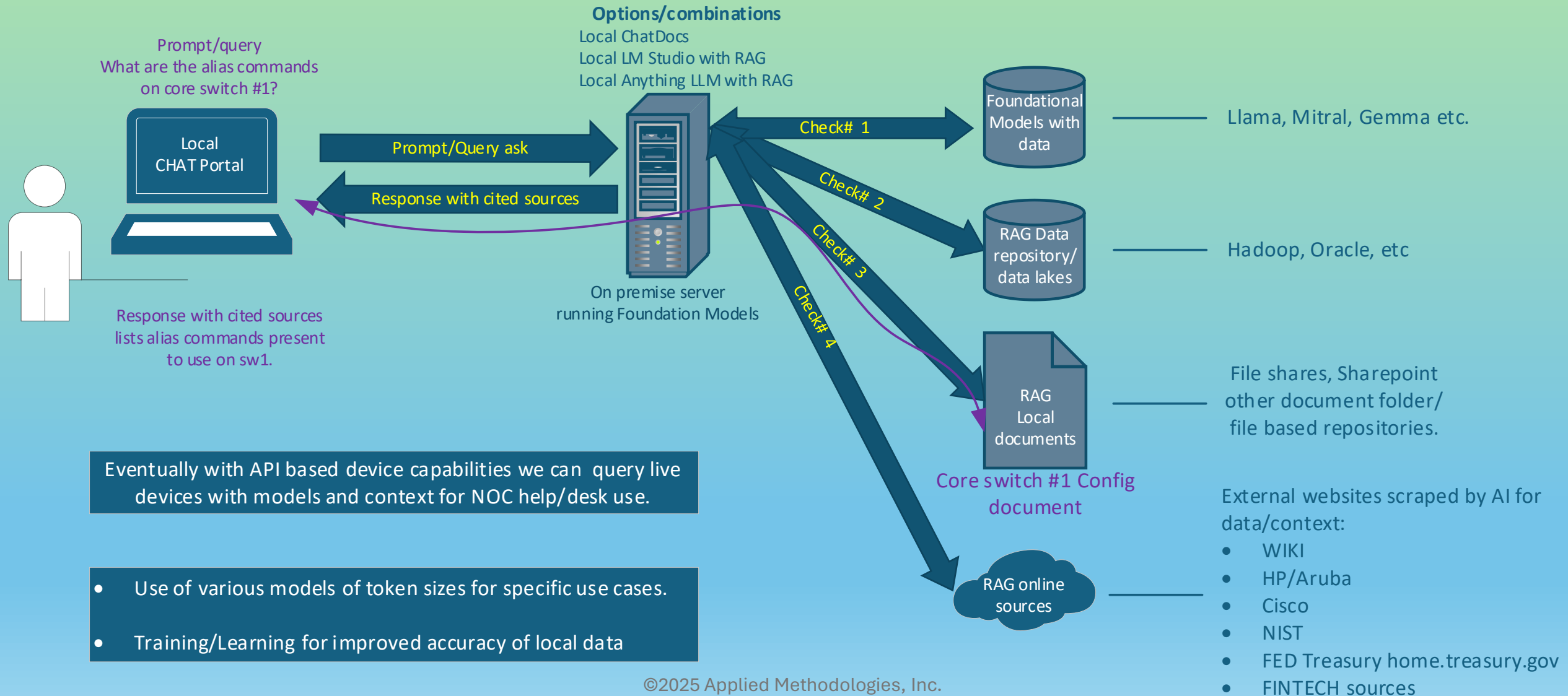
- Secure leadership commitment
- Take stock of their data ecosystem
- Unlock legacy data
- Make data accessible to all
- Promote self-service
- Provide continued education
- Shift company culture
- Change the way they store and distribute data
- Data democratization can help organizations become more data-driven. It can also help employees, stakeholders, and consumers feel confident working with data and trust the end results.
- Some synonyms for data democratization include: data openness, data accountability, data availability, and data engagement.

Context document sources

Some examples include:

- Technical and business policies
- Technical and business procedures
- Compliance/governance documents
- Log files for quick troubleshooting or security incident querying
- Help-desk guides
- IT device configuration files
- IT product documentation
- Customer account documents for reports or sales use

General Architectural Framework



Initial first step into AI tools - ChatDoc

- Can be used in their cloud AlaaS for “scrubbed” data.
- Can be used in their cloud AlaaS service for compliant data.
- Can be used locally with no connectivity to internet.
- Simple tool to get started using AI to “chat with your documents”

https://chatdoc.com/

ChatDOC

PRODUCT HUNT
#1 Product of the Day

ChatPaper

API

Pricing

Blog

Changelog

Get Started

NEW Chat seamlessly with PDFs using our Chrome Extension →

Chat with documents. Get instant answers with cited sources.

Dive into PDFs like never before with ChatDOC. Let AI summarize long documents, explain complex concepts, and find key information in seconds.

→ Try for Free

Supported Formats

PDF

DOC

DOCX

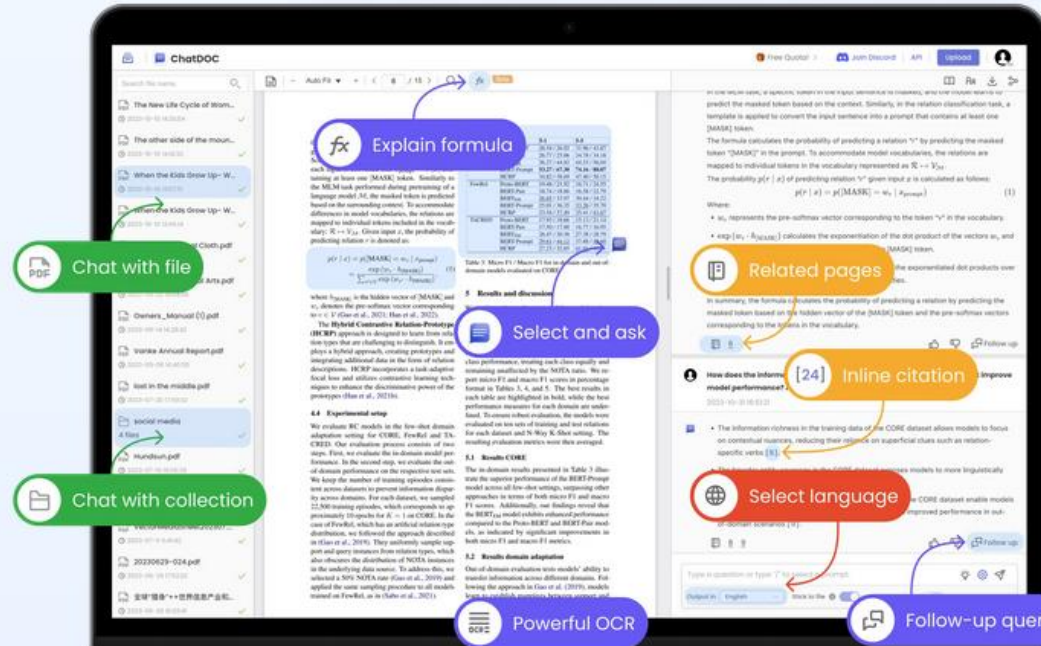
SCAN

WEBSITE

EPUB

MD

TXT



ChatDoc How it works

1. Log into site – local or cloud
2. Drag your documents you want to chat with
3. Start chatting with your documents

Drag a PDF or other document into ChatDoc

Example uses a PDF from NetworkWorld Linux training series

The screenshot shows the ChatDoc web application interface. The browser address bar displays `https://chatdoc.com/chatdoc/#/upload`. The interface includes a sidebar with a search bar and a list of files. The main area features two large buttons: "Click or drag and drop to upload." and "Drag and drop folder as collection." Below these buttons are links for "Upload files through URL" and "Chat with website". A "Support:" section lists various file formats: pdf, doc, docx, markdown, epub, txt, website, and OCR. A comparison table between "Free" and "Pro" plans is also visible.

Free	Pro
20 pages/pdf	No PDF page limit
OCR	500 OCR pages/30 days
doc, docx, markdown, epub, txt, website	300K tokens/doc, docx, markdown, epub, txt, website
36 MB /file	36 MB /file
2 files/24 hours	300 files/30 days
30 files/collection	No file limit/collection
20 questions/24 hours	300 questions/24 hours
GPT-4o	GPT-4o (extra paid)
	Unlock image analysis and enhanced responses!

The Windows File Explorer window at the bottom shows the path `C:\Users\jsicu\OneDrive - App > Jeff - Applied Methodologies, Inc > Documents > AI TOOL Context repository`. A file named "The Linux command-line cheat sheet..." is selected, and a hand-drawn arrow points from it to the "Click or drag and drop to upload." button in the ChatDoc interface.

Start chatting with your documents

Already provides sample query/prompts to use based on context of document

The screenshot displays the ChatDOC web application. The main window shows a document titled "The Linux command-line che..." with a table of Linux commands. The sidebar on the right contains a "Recommend" section with sample prompts related to the document content.

Command	Function	Example
pwd	Displays your current location in the file system	pwd
whoami	Displays your username – most useful if you switch users with su and need to be reminded what account you're using currently	whoami
ls	Provides a file listing. With -a , it also displays files with names starting with a period (e.g., .bashrc). With -l , it also displays file permissions, sizes and last updated date/time.	ls ls -a ls -l
env	Displays your user environment settings (e.g., search path, history size, home directory, etc.)	env
echo	Repeats the text you provide or displays the value of some variable	echo hello echo \$PATH
history	Lists previously issued commands	history history tail -5
passwd	Changes your password. Note that complexity requirements may be enforced.	passwd history tail -5

Examining files

Linux provides several commands for looking at the content and nature of files. These are some of the most useful commands.

ADVERTISING

Recommend Custom

1. What are some essential Linux commands included in the cheat sheet? →
2. How can Linux commands assist users in learning about other commands? →
3. What role does privileged access play in shutting down or rebooting Linux systems? →
4. How does the choice of text editor impact the process of creating files and directories in Linux? →

What resources are available for users to explore additional Linux commands related to

Type a question or type "/" to select a prompt.

gpt-4o-mini English Stick to file

Pulls contextual data from document and uses LLM model for additional information.

ChatDOC

Free Quota > | ChatPaper | Browser Extension | API | Upload

Search file name.

Created ▾ | All files →

The Linux command-line che...
2025-01-02 12:15:56
No More

Auto Fit

< 2 / 13 >

10 Leadership

NETWORKWORLD

what does the ls command provide?

2025-01-02 12:22:14

The `ls` command provides a file listing in the Linux operating system. It can be used with various options to modify its output. For instance, using `-a` displays files, including those with names starting with a period (hidden files), and using `-l` provides a long format that includes file permissions, sizes, and the last updated date/time. An example of its usage is simply entering `ls`, or `ls -a` and `ls -l` for more detailed listings.

2 5 4 1

Follow up

Type a question or type "/" to select a prompt.

gpt-4o-mini | English | Stick to file

Command	Function	Example
pwd	Displays your current location in the file system	pwd
whoami	Displays your username – most useful if you switch users with <code>su</code> and need to be reminded what account you're using currently	whoami
ls	Provides a file listing. With <code>-a</code> , it also displays files with names starting with a period (e.g., <code>.bashrc</code>). With <code>-l</code> , it also displays file permissions, sizes and last updated date/time.	ls ls -a ls -l
env	Displays your user environment settings (e.g., search path, history size, home directory, etc.)	env
echo	Repeats the text you provide or displays the value of some variable	echo hello echo \$PATH
history	Lists previously issued commands	history history tail -5
passwd	Changes your password. Note that complexity requirements may be enforced.	passwd history tail -5

Examining files

Linux provides several commands for looking at the content and nature of files. These are some of the most useful commands.

Query with answer and cited pages to refer to shown

ChatDOC

Free Quota

ChatPaper

Browser Extension

API

Upload

Search file name.

Created

First Bank of LI Port Security H...

No More

Auto Fit

11 / 19

what are the steps to disable port security?

2025-01-02 12:30:14

To disable port security on a port, the following steps can be followed as outlined in the document.

4o-mini

1. Log into the Switch GUI.

2. Navigate the menu to the path: Switching -> Network Security -> Port Security[12].

3. Locate the port that you wish to disable port security for. Ensure that the user's original authorized device is plugged back in before proceeding[10a].

4. In the settings for that specific port, change the status from "Locked" to "Unlocked," indicating that port security is not applied (turned off) on the port[10b].

5. Click the "Apply" button to save the changes[10a].

This process will effectively disable port security for the selected port, allowing any device to connect without restrictions.

12 10

Follow up

5 MOVES ADD CHANGES(MAC)

For handling MACs IT will need to unlock a port or multiple ports for a deployment of new or changed IT devices connecting to the network, thus the following steps are outlined.

1. You log into the Switch GUI and navigate the menu to path Switching->Network Security -> Port security as depicted in yellow highlight below.

2. The current user's or unused port should be in a locked state as depicted in yellow highlight below.

PORT SECURITY ADMINISTRATION

Deployment note:
After IT deploys the Port Security configuration to all or most of the HQ switch ports an easier way to determine if the port has been configured for port-security is to use the following GUI page as depicted below, then toggle the Set Port column to sort on ports that are in a Locked state.

Remember Locked indicator in the Set Port column means that port security is applied, not that a violation is underway.

Port Security Administration

Port Security Show All

Port Settings

Port	Set Port	State	Port Frequency	Security Mode	State
1	Locked	Locked	30	Secure	CS
2	Locked	Locked	30	Secure	CS
3	Locked	Locked	30	Secure	CS
4	Locked	Locked	30	Secure	CS
5	Locked	Locked	30	Secure	CS
6	Locked	Locked	30	Secure	CS
7	Locked	Locked	30	Secure	CS
8	Locked	Locked	30	Secure	CS
9	Locked	Locked	30	Secure	CS
10	Locked	Locked	30	Secure	CS







©2022 APPLIED METHODOLOGIES, INC.

Next Step - build your private Local LLM with RAG

- Utilize many existing Foundation Models(FM) in place today.
- See Hugging Face for many open-source free models to use, train and test
<https://huggingface.co/>
- Combine with your document repository for contextual and insightfulness chatting with your documents.
- Utilize RAG for local and external sites.
- Example: Chat with a local document on a specific internal policy and it utilizes RAG to validate context against an external authoritative web site.


LM STUDIO + ANYTHING LLM

Discover, download, and run local LLMs

-  • Run LLMs on your laptop and DC servers, entirely offline
-  • Chat with your local documents(Laptop use) Scale for organization on server for all shared use local documents.
-  • Use models through the in-app Chat UI or an OpenAI compatible local server.
-  • Download any compatible model files from **Hugging Face**  repositories
-  • Discover new & noteworthy LLMs right inside the app's Discover page
- Use as a back end LLM server for use with your own API based portal or another portal – Anything LLM.

https://lmstudio.ai/

Can use as “back end” and primary server

 **LM Studio**

Discover, download, and run local LLMs

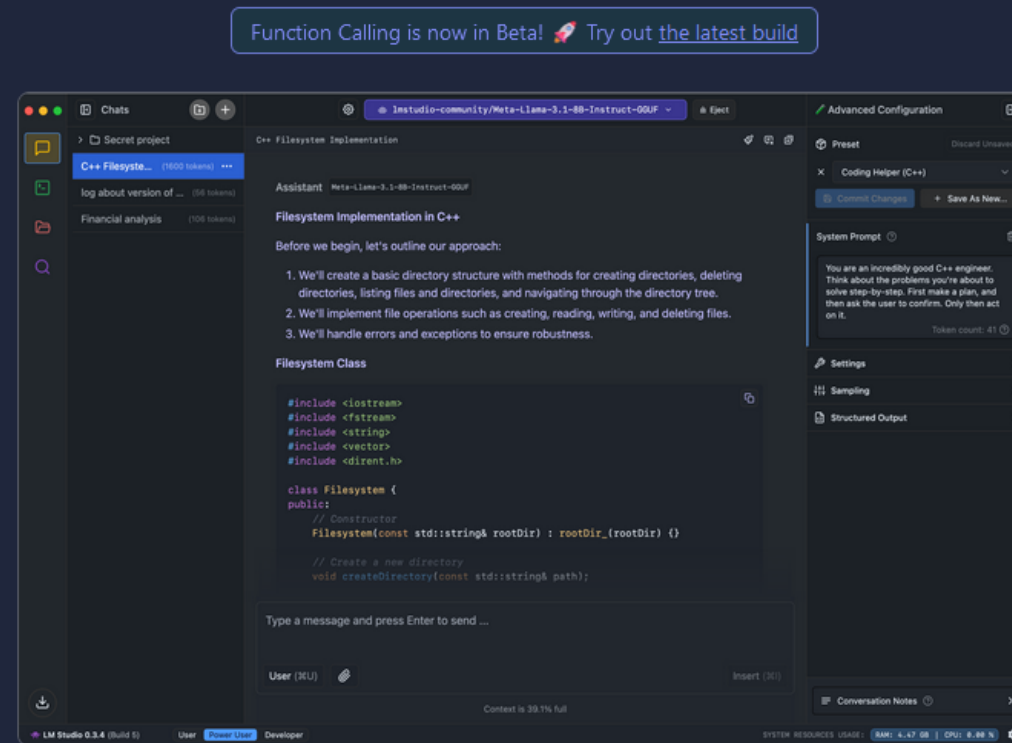
Run Llama 3.2 Mistral Phi Gemma DeepSeek Qwen 2.5 on your computer ⓘ

 Download LM Studio for Mac (M series) 0.3.5

 Download LM Studio for Windows 0.3.5

 Download LM Studio for Linux 0.3.5

LM Studio is provided under the [terms of use](#)



<https://anythingllm.com/>

AnythingLLM


Products ▾ Resources ▾ Community ▾

The all-in-one AI application

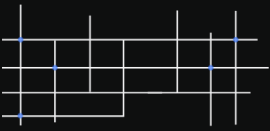
Everything great about AI in one desktop application.
Chat with docs, use AI Agents, and more - full locally and offline.

[Download for desktop](#)

[Star on Github](#)




Leverage powerful AI tooling with no set up



Supports custom models

Stay local fully with our built-in LLM provider running any model you want.


Or leverage enterprise models from OpenAI, Azure, AWS, and more.



Use every type of document.

PDFs, word documents, CSV, codebases, and so much more make up your business - now you can use them all.

Even import docs from online locations



Full privacy

AnythingLLM comes with sensible and locally running defaults for your LLM, embedder, vector database, and storage and agents.


Nothing is shared unless you allow it.

Tons of features built-in

AnythingLLM ships with a ton of benefits and features built-in without writing a single line of code.

[Download for desktop](#)


[Get Started on Cloud](#)



Any LLM anytime


Run your preferred LLM locally in a click, or connects to your local or cloud LLM engine.

All with no frustrating set up.




Simple Interface

AnythingLLM wraps all of this incredible functionality into a sleek UI for you to leverage LLMs easily for any task.



Open Source


AnythingLLM is open source and free to use & MIT licensed.



Customizable

Extend AnythingLLMs with custom agents, data connectors, and more.


With your help, there is nothing AnythingLLM cannot do.



Multi-model


AnythingLLM works with both text only and multi-modal LLMs in a single interface.

Can with images or audio easily.




Built-in Developer API

AnythingLLM is also suitable to be used as a powerful API for any custom development or feature in existing products.



Ecosystem

A growing ecosystem of plugins and integrations to extend its functionality across the app and your organization.



Privacy-focused

Designed to be private by default. Everything is stored and run locally on the machine running AnythingLLM.

©2025 Applied Methodologies, Inc.

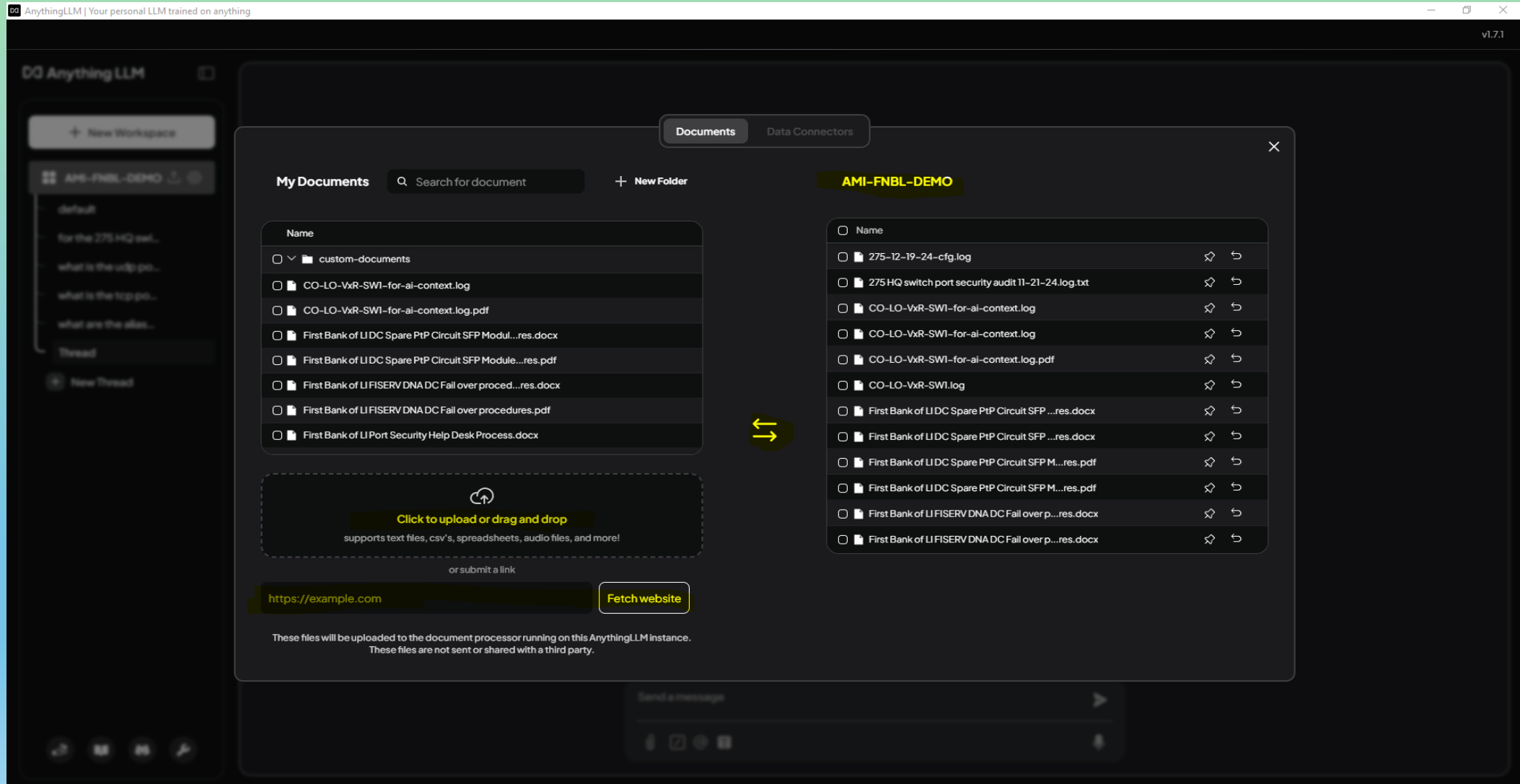
LM Studio – Local Server can use stand alone with any model or your documents

The screenshot displays the LM Studio 0.3.5 application window. The top bar shows the title "LM Studio - 0.3.5" and a "Select a model to load (Ctrl + L)" dropdown. The main interface is divided into three sections:

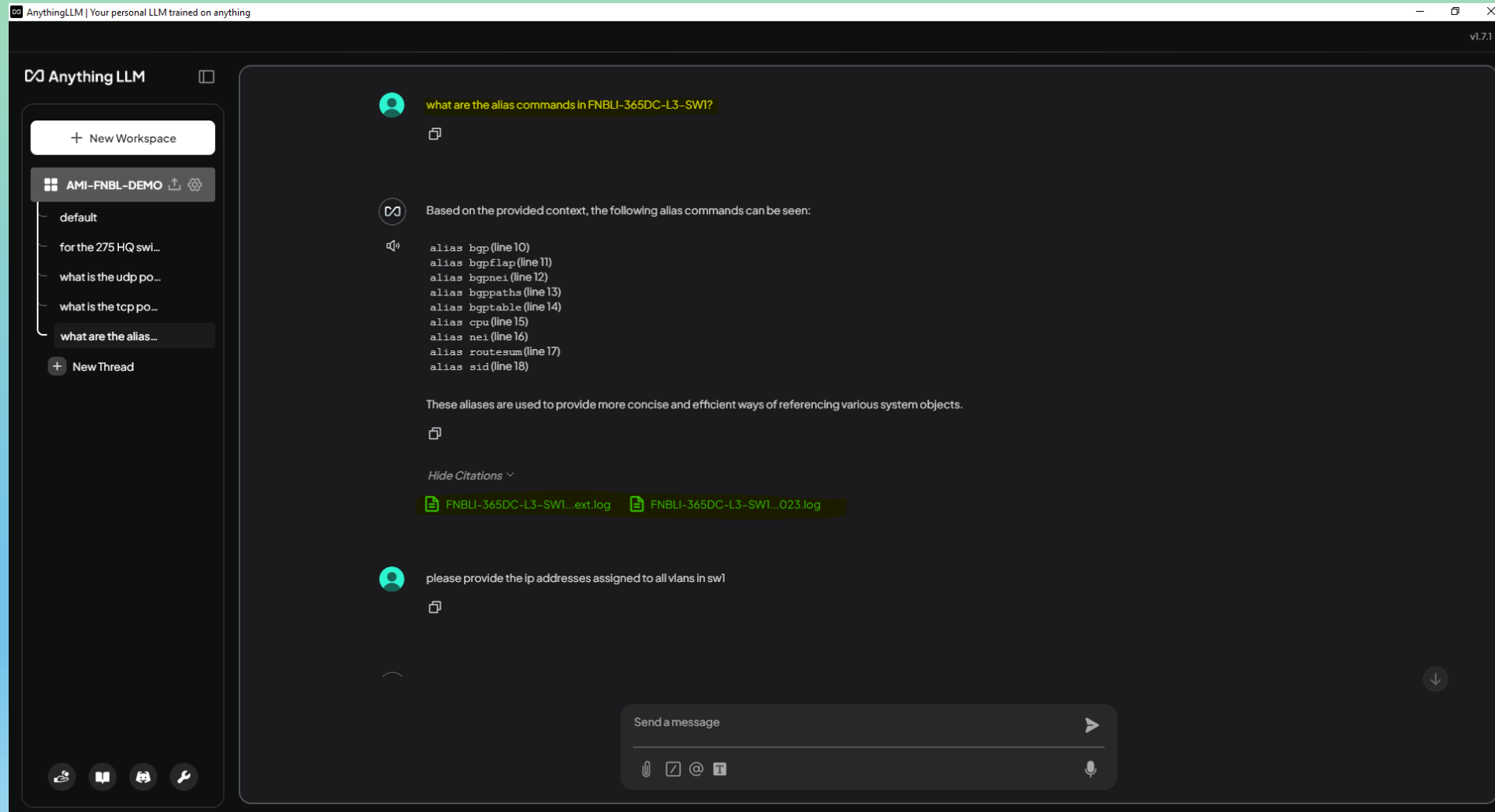
- Left Sidebar:** Contains navigation icons for chat, server status, file explorer, and search. The "Local Server" tab is highlighted.
- Central Panel:**
 - Server status:** Shows "Running" with a "Stop Server (Ctrl + J)" button. Below are settings for "Server Port" (1234), "Enable CORS" (disabled), and "Serve on Local Network" (disabled).
 - Supported endpoints (OpenAI-like):** A dropdown menu showing endpoints like "GET /v1/models", "POST /v1/chat/completions", "POST /v1/completions", and "POST /v1/embeddings".
 - Loaded models:** A list showing a loaded model "llm llama-3.2-1b-instruct" with a size of "1.32 GB" and an "Eject" button.
 - Developer Logs:** A log window showing server startup messages and API requests.
- Right Panel:**
 - Model Information:** Displays details for the loaded model, including "Model" (hugging-quants/Llama-3.2-1B-Instruct-Q8_0-GGUF), "File" (llama-3.2-1b-instruct-q8_0.gguf), "Format" (GGUF), "Quantization" (Q8_0), "Arch" (llama), "Domain" (llm), and "Size on disk" (1.32 GB).
 - API Usage:** Shows the model's API identifier and a confirmation message: "The local server is reachable at this address" with the URL "http://127.0.0.1:1234".

The bottom status bar indicates "LM Studio 0.3.5 (Build 2)" and "SYSTEM RESOURCES USAGE: RAM: 2.04 GB | CPU: 0.00 %".

Anything LLM works as front end to LM Studio - can use additional models and add documents and web links for RAG



Sample query using Anything LLM as front end and Local LM as back end – note the citations



Another local query using the model and local documents for context

The screenshot displays the AnythingLLM application interface. The top bar shows the title 'AnythingLLM | Your personal LLM trained on anything' and the version 'v1.7.1'. The left sidebar contains a 'New Workspace' button, a workspace named 'AMI-FNBL-DEMO', and a list of threads including 'default', 'for the 275 HQ swi...', 'what is the udp po...', 'what is the tcp po...', and 'what are the alias...'. A 'New Thread' button is also present. The main area shows a query: 'please provide the ip addresses assigned to all vlans in sw1'. The response, generated by the model, states: 'Based on the provided context, the IP addresses for all VLANs in SW1 (Server) are:'. It then lists five VLANs with their IP addresses and subnets: VLAN 10: 10.203.10.0/24, VLAN 100: 10.100.100.0/29, VLAN 100: 10.100.100.0/29, VLAN 100: 10.203.10.24 (no VRRP mode active-active), and VLAN 200: 10.203.10.24/24. A note mentions the IPv6 address is [REDACTED]. At the bottom, three source documents are listed: 'CO-LO-VxR-SW1-for-...log.pdf', 'CO-LO-VxR-SW1.log', and 'CO-LO-VxR-SW1-for-...ext.log'. The bottom of the interface features a 'Send a message' input field with icons for attachments, links, mentions, and a microphone.

AnythingLLM | Your personal LLM trained on anything v1.7.1

AnythingLLM

+ New Workspace

AMI-FNBL-DEMO

- default
- for the 275 HQ swi...
- what is the udp po...
- what is the tcp po...
- what are the alias...

+ New Thread

alias routesum (line 17)
alias sid (line 18)

These aliases are used to provide more concise and efficient ways of referencing various system objects.

Show Citations >

please provide the ip addresses assigned to all vlans in sw1

Based on the provided context, the IP addresses for all VLANs in SW1 (Server) are:

- VLAN 10: 10.203.10.0/24
- VLAN 100: 10.100.100.0/29
- VLAN 100: 10.100.100.0/29
- VLAN 100: 10.203.10.24 (no VRRP mode active-active)
- VLAN 200: 10.203.10.24/24

Note that the IPv6 address is [REDACTED]

Hide Citations <

- CO-LO-VxR-SW1-for-...log.pdf
- CO-LO-VxR-SW1.log
- CO-LO-VxR-SW1-for-...ext.log

Send a message

Side by side of both tools working together

The image shows a Windows desktop with two applications open. On the left is 'AnythingLLM', which has a dark theme. It features a sidebar with a '+ New Workspace' button and a list of workspaces including 'AMI-FNBL-DEMO'. The main chat area shows a conversation where a user asks for alias commands on a specific switch, and the AI provides a list of commands. On the right is 'LM Studio', which has a light blue header. It shows the server status as 'Running' and a list of loaded models, including 'llm llama-3.2-1b-instruct'. It also displays developer logs and model information on the right side.

Combined tools working together

The image displays two web applications side-by-side, demonstrating their combined use for running Large Language Models (LLMs).

AnythingLLM (Left Panel):

- Header:** AnythingLLM | Your personal LLM trained on anything
- Left Sidebar:** Includes a "New Workspace" button, a workspace named "AMI-FNBL-DEMO", and a list of threads. The selected thread is "what are the alias...".
- Main Chat Area:** Shows a conversation with a user asking "What are the alias commands on FNBLI-365DC-L3-SW1". The assistant responds with a list of alias commands based on provided context.

```
alias bgp (line 10)
alias bgpflap (line 11)
alias bgpnei (line 12)
alias bgppaths (line 13)
alias bgptable (line 14)
alias cpu (line 15)
alias nei (line 16)
alias routesum (line 17)
```
- Bottom Bar:** Includes a "Send a message" input field and icons for attachments, links, mentions, and emojis.

LM Studio (Right Panel):

- Header:** LM Studio - 0.3.5
- Left Sidebar:** Includes buttons for "Server status", "Supported endpoints", and "Developer Logs".
- Main Chat Area:** Displays the JSON-RPC messages exchanged between the client and the server, showing the chat completion request and response.

```
{
  "role": "user",
  "content": "What are the alias commands on FNBLI-365DC-L3-SW1?"
},
{
  "role": "assistant",
  "content": "Based on the provided context, the following alias commands can be seen:\n\n`alias bgp` (line 10)\n`alias bgpflap` (line 11)\n`alias bgpnei` (line 12)\n`alias bgppaths` (line 13)\n`alias bgptable` (line 14)\n`alias cpu` (line 15)\n`alias nei` (line 16)\n`alias routesum` (line 17)\n\nThese aliases are used to provide more concise and efficient way s of referencing various system objects."
},
{
  "role": "user",
  "content": "please provide the ip addresses assigned to all vlans in sw1"
},
{
  "role": "assistant",
  "content": "Based on the provided context, the IP addresses for all VLANs i n SW1 (Server) are:\n\nVLAN 250: `10.203.250.230/24`\nVLAN 1700: `10.170.100.1/29`\nVLAN 1900: `10.190.100.3/29`\nVLAN 2000: `10.203.100.230/24` (no VRRP mode active-active)\nVLAN 2001: `10.203.101.230/24`\n\nNote that the IPv6 addre ss is `10.203.254.1/32`."
},
{
  "role": "user",
  "content": "What are the alias commands on FNBLI-365DC-L3-SW1"
}
],
"temperature": 0.7
}
```
- Right Sidebar:** Contains "Model Information" and "API Usage" sections. The model is "hugging-quants/Llama-3.2-1B-Instruct-q8_0-GGUF". The API usage section shows the model's identifier and a checkmark indicating the local server is reachable at the address "http://127.0.0.1:1234".
- Bottom Bar:** Includes "LM Studio 0.3.5 (Build 2)", user roles (User, Power User, Developer), and system resource usage (RAM: 2.06 GB, CPU: 0.00 %).

Footer: ©2025 Applied Methodologies, Inc. 1:58 PM 1/2/2025

Learning / Training

- You can leave your setup in a static data state or have it learn/train based on your local data, old or new, for improved results and to achieve detailed chatting results.
- You can add different models from Huggingface or vendors for specific uses such as network protocol analysis, network device behavior from a vendor.
- Learning/Training is dependent on your setup, model(s) and tools used.

Chatting with live “production” devices(future)

- Chat with your devices via your equipment vendor’s API.
- Chat with your networking, server/compute and other IT infrastructure devices in real time for troubleshooting and operational behavior status.
- Network/Server/Application/Cloud Vendors are already integrating this capability into their next generation of applications, management tools and equipment.

Summary

- Democratize your local data in a secure local on-premise environment.
- Provides staff ability to use documents in a fluid manner to achieve higher daily productivity.
- Easy to implement and manage with current generation of AI local LLM tools today.
- Back end tools such as LM Studio, Anything LLM, GPT4All, Ollama, can all be used locally.
- Can be used as secure on premise AI first step for organizations to embrace AI for internal domain use.
- Can be combined or migrated to sanctioned Microsoft or Google company used AI solutions when organization formally embraces AI.
- Tools are always improving in this exciting technology for the benefit of the organization.