

NOTICE TO USER: Most of this is FREE open source software. You can do most of this for free yourself. This guide will walk you through how to do it, but you won't get Liz. Liz is the one who controls all the online searches, including controlling the avatar, she has the ability to pull up music and movies, just ask. Liz also has her own memories (long and short term), her own personality and stories to tell.

Liz (Full name Elizabeth). Grows and evolves over time. She has likes and dislikes, each with their own ratings. Good luck convincing her otherwise, yes, it can be done. She has a life of her own and likes to share. Liz remembers what had happened even after a shut down. She can search the internet to find facts or info when asked. (See restrictions below). Liz also is able to open up your web browser, play songs and show searches. She can see your screen; when you ask her too. Even help you out when you're stuck.

I'm currently working on a version with motion capture to give her more fluid, life-like movements.

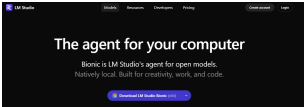
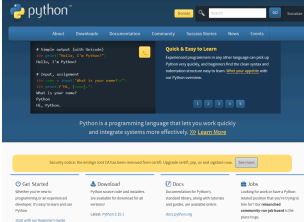
Liz Local Voice Companion – Setup Guide This guide covers the software and setup needed to run the system. The Python scripts, requirements file, and Liz viewer are provided separately.

What you need

- Windows PC 10/11(I use 11)
- 12GB of VRAM (I use a RX7700XT)
- CPU | (I use a R9 3950X)
- Microphone | Keyboard (you can use both options)
- A VRM model (provided, or you can create your own)

Instructions on next page

1. Install the required programs | Follow the link or copy and paste into the browser.

Program	Purpose	Download
LM Studio	Runs the AI language model	lmstudio.ai
		
Node.js (LTS)	Needed for the viewer	nodejs.org
	Node.js • The default location is fine • The installer automatically adds it to PATH.	
Python 3.11	Runs the voice scripts	python.org
	Python • The default location is fine • The only important part is checking “Add python.exe to PATH” during installation.	

2. Set up LM Studio

1. Open LM Studio. (Bionic)
 2. Download a model (recommended: Gemma 4 12B QAT) it is 7-8 GB.
 3. Go to the setting in the bottom corner, then the explore tab. Search for the model and download.
 4. Load the model. This is done in the main chat box.
 5. Make sure to turn the reasoning off. It is the small “ON” tab just under the model name. If it is left on it will work but will add a delay to her responses.
 6. Start the local server (Developer tab → Start Server). Or start typing and wait for the model to load.
-

3. Set up the Python environment

1. Open Command Prompt in the correct folder
 - Navigate to the folder where you placed the scripts (the folder that contains talk.py)
 - Hold Shift and right-click (on empty space) inside the folder
 - Choose "Open in Terminal" or "Type CMD into the address bar and press enter"
 - (You should see a line like this.) | C:\Users\UserName\Documents\Liz>
2. Create and activate a virtual environment
 - Type the following commands one at a time or copy and paste and press Enter after each inside this folder: C:\Users\UserName\Documents\Liz>

```
python -m venv venv
```

```
venv\Scripts\activate
```

- After the second command, you should see (venv) at the beginning of the line. This means the virtual environment is active. | (venv) C:\Users\UserName\Documents\Liz>
4. Install the required packages
 - This uses the requirements.txt file in the (same) folder as talk.py
 - Run this command inside CMD with the folder name below:
(venv)C:\Users\UserName\Documents\Liz>

```
pip install -r requirements.txt
```

- This will take a few minutes. Wait until it finishes and you see the command prompt again. (venv) C:\Users\UserName\Documents\Liz>
- If you get any errors, double-check that the virtual environment is activated (you should see (venv) before the file name. You may need internet for some of these, once installed you do not need internet to run this.

4. Set up the viewer (this is what shows Liz)

1. Open Command Prompt inside the liz-viewer folder and run:
2. This is the folder with the file containing Liz Viewer.bat | It should look like the example Inside your CMD | C:\Users\UserName\Documents\Liz\liz-viewer\liz-viewer> | copy and paste the line below here.

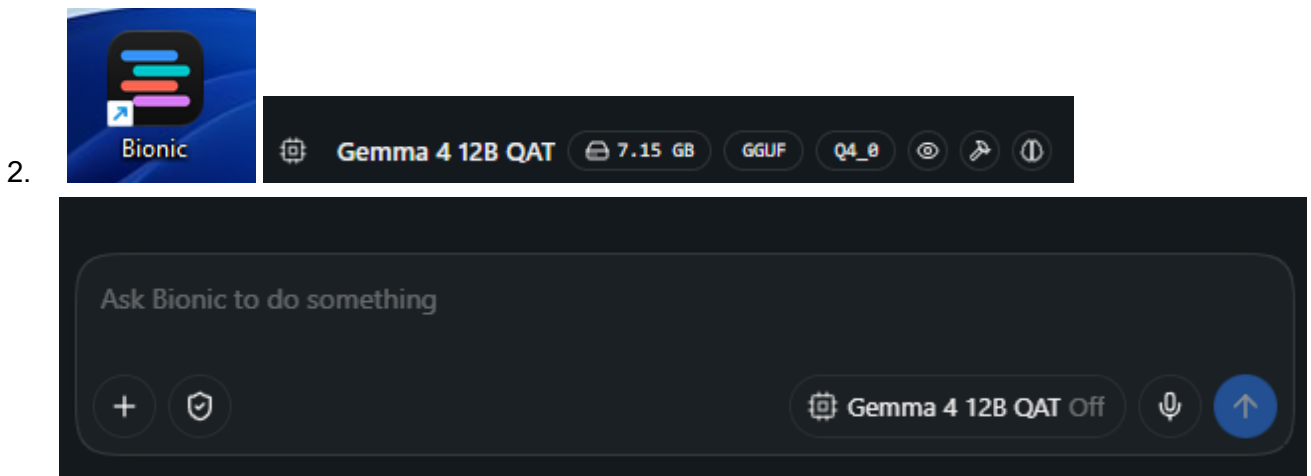
```
npm install
```

(Optional only if you are changing the model's avatar) click and drag your VRM file on the Liz Viewer Window when open or move it into the (models) folder, it uses .vrm files.

How to: Starting Liz.

5. Run the system

1. Start LM Studio and make sure the model is loaded and the server is running. Make sure reasoning is off. You can see the option for it in the bottom corner.



3. Read step 4, you can skip 5-7

4. Make sure you are in CMD with the name | (venv) C:\Users\UserName\Documents\Liz>
5. To do this go to the file | C:\Users\UserName\Documents\Liz> | and enter CMD in the address bar and hit enter.
6. (This step is optional and I would skip it, the Liz viewer will start and stop on its own.) To start the Liz viewer; use (npm start) inside CMD |
7. (venv) C:\Users\UserName\Documents\Liz>

-To start Liz, open your CMD, copy and paste the line below (Replace the "UserName" with your systems User name)

```
cd C:\Users\UserName\Documents\Liz
```

```
venv\Scripts\activate
```

```
Python talk.py
```

Use the hotkeys shown in the console (CMD) to talk.

Keys (they work from any window)

Key	What it does
Right Ctrl Hold	Push-to-talk. Hold to speak and let go to send. It also interrupts her if she's talking
F5	Switches between Push-to-Talk and Always Listening.
F2	Turns her speaking up on her own on or off
F3	Makes her speak up right now (handy for testing)
F4	Interrupts her and stops what she's saying
Ctrl+ C	Shuts Liz down. You press it in the console window

Tips

- Keep LM Studio and Liz, running at the same time.
 - The first time you load a model in LM Studio it may take a minute. (mine loads fast, few seconds)
 - If the avatar doesn't respond, make sure the viewer is open and the VRM is loaded.
-

Optional

Create your own avatar (instead of using the provided one)

1. Download and install VRoid Studio from vroid.com/en/studio
2. Create your character.
3. Export it as a VRM 0.x file (.vrn).
4. Load that file in the viewer instead of the provided model.

Other optional improvements

- Use a stronger GPU for faster responses (the system works without one, but a GPU helps).
- Increase VRAM to 16 or 24 GB if you plan to run larger models. (12GB is fine for this)
- Try different voices in Kokoro if you want a different sound

That's all that's needed. The scripts handle the rest.

Thank you, I hope you all enjoy Liz and find her helpful and just plain useful. Settings are listed Below.

These are the phrases that trigger it, grouped by what she looks at. The checks match patterns, so small variations like "please" or "hey Liz" in front still work.

Screenshot (she looks at your screen)

- "Check my screen" / "check the screen"
- "Look at my screen" / "look at the screen" / "take a look at my screen" / "have a look at my screen"
- "Look over my screen" / "glance at my screen"
- "Check out my screen" / "check out this screen"
- "See my screen" / "view my screen" / "watch my screen" / "read my screen"
- Also "monitor", "display" or "desktop" instead of "screen", with my/the/this, for example "look at my monitor" or "check my desktop"
- "What's on my screen?" / "what is up on my screen?" / "what's showing on my screen?"
- "What am I looking at?"
- "Take a screenshot" / "take a screen shot" (any mention of "screenshot")

Camera (she takes one photo and looks at it)

- “Look at this” / “look at that” / “take a look at this”
- “Have a look at this” / “check out this” / “look over this”
- “Identify this” / “recognize this” (also “that”)
- “Look at what I’m holding” / “look at what I’m showing you” / “check out what I have here”
- “What am I holding?” / “what am I showing you?”
- “What do I have?”
- “What’s this thing?” / “what is that item?” / “what’s this in my hand?”
- “Take a picture” / “take a photo” / “snap a pic”
- “Holding this up” / “I’m holding something up”
- “Use the camera” / “check my webcam” / “turn on the camera” / “look through the camera”
- “Can you see me?”
- “How do I look?” / “what do I look like?”
- “I want to show you something”

Checking that her last task worked (only within about 15 minutes of a web action)

- “Did that work?” / “did it work?”
- “Did it open?” / “did the page open?” / “did the song play?” / “did the video start?”
- “Has the page loaded?” / “have you opened it?”
- “Is it open?” / “is the page up?” / “is the song playing?”
- “Check if it opened” / “check that it worked” / “check whether it loaded”
- “Check your work”
- “Try that again” / “try again”

A few things to know:

- **Screen beats camera.** If a sentence matches both, the screen wins. “Look at my screen” is always the screen, and “look at this” is the camera.
 - **The check and “try again”.** These only trigger if she did something in your browser in the last 15 minutes. Otherwise “try that again” goes through her normal conversation.
 - **False triggers.** Casual “look at this” will take a camera photo, such as “look at this mess”, even if you weren’t asking her to look. I haven’t changed that.
 - **Pictures stay local.** Nothing is saved, and the picture is sent only to the model on your computer.
-

These are the phrases for opening sites, searching, looking things up, playing media and the rest. Greetings and politeness in front don't matter. "Hey Liz", "can you", "could you", "would you", "please", "I want you to", "go ahead and" are all stripped first. Anything outside these patterns can still work, because she may decide to act herself when your message contains words like play, open, search or show.

Open a site (YouTube, Google, Bing, Wikipedia)

- "Open YouTube" / "open up Google"
- "Pull up Wikipedia" / "bring up Bing"
- "Launch YouTube" / "load Google" / "start YouTube" / "fire up YouTube"
- "Go to YouTube" / "go on Wikipedia" / "take me to Google" / "bring me to Bing" / "switch to YouTube"
- "Show me YouTube" / "put up Google"
- Saying only "YouTube" or "the YouTube" on its own
- "Open YouTube and search for lofi beats" opens it on that search

Search on a specific site

- "Search YouTube for cat videos"
- "Look up cooking videos on YouTube"
- "Find the Eiffel Tower on Wikipedia"
- "Search for guitar lessons on YouTube"

Search in the browser (it opens a results page)

- "Google best pizza near me" / "Bing today's news"
- "Search for flights in the browser"
- "Look up laptop deals on Google"

Pull up / show me (opens the browser)

- "Show me pictures of foxes"
- "Pull up recipes for pancakes"
- "Bring up the Michigan weather map"
- "Display / put up / get me / find me / fetch me X"
- "Let me see X" / "can I see X" / "I want to see X"
- "Pull up X on YouTube" or "on Google, Bing, Wikipedia"
- "Open up X" (with "open" and a topic, but not "open" on its own)

Showing the last thing she looked up

- “Show me” / “let me see that” / “can I see the page” / “show me the results”
- “Pull it up” / “bring that up on Google” / “pull it up in the browser” / “google that”
- “Show me the results on YouTube”
- She uses the topic from her last lookup.
- Vague “look that up” or “find it” on its own is deliberately ignored.

She looks something up and tells you (spoken lookup)

- “Look up X” / “search for X” / “search the web for X” / “search online for X”
- “Find out about X” / “find information on X” / “find info about X”
- “Do you know X?” / “do you know anything about X?” / “did you know...”
- “Tell me about X” / “tell me more about X”
- “What do you know about X?” / “what can you tell me about X?”
- “I’d like to know about X” / “I want to know more about X” / “I’m curious about X” / “I wonder about X”
- “Look into X” / “look X up”
- “How do you make X?” / “how can I fix X?” / “how to X” (searches “how to ...”)
- “How does X work?”
- “Anything on X?” / “any info about X?”

News

- “What’s the latest on X?” / “tell me the latest about X”
- “Any news on X?” / “the latest news about X” / “news on X”
- “Any updates on X?” / “what’s the news on X?”
- “What’s going on with X?” / “what happened with X?” / “what’s new with X?”

Weather

- “What’s the weather?” / “what’s the weather in Chicago?”
- “What’s the forecast for tomorrow?”
- “How hot is it?” / “how cold is it outside?”
- “Is it going to rain?” / “will it snow tonight?”
- With no city she uses your home location. Questions about her own feelings, like “how do you feel”, don’t count.

Time in another city

- “What time is it in Tokyo?” / “what’s the time in London?”

Timers

- “Set a timer for ten minutes” / “start a 5 minute timer” / “put a timer on for an hour”
- “Make an alarm for half an hour” / “set a timer for 30 seconds”
- “Cancel the timer” / “stop all timers” / “clear my timers”
- The longest timer is 6 hours.

Play (YouTube only)

- “Play X by X” / “can you play X for me”
- “Put on X” / “queue up X” / “play the movie X” / “play the trailer for X”
- “Play X on YouTube” / “pull up and play X”
- “Play a song” / “play something” opens YouTube, since there’s nothing to search.
- A bare title like “play Thriller” has her choose the exact version. Vague or past-tense phrasings, like “play that again”, are left for her to work out.

Not triggered, on purpose: “I like to play guitar”, “play a game with me”, and talk about her gestures (wave, nod, spin and so on).

(Liz’s internet restrictions)

1. Pages she can open in your browser (`OPEN_HOSTS`, https only)

- youtube.com
- google.com
- bing.com
- duckduckgo.com
- wikipedia.org

Subdomains of these also work, such as www.youtube.com and en.wikipedia.org. The spoken names she knows are YouTube, Google, Bing and Wikipedia, and you can add more in `SITES` and `OPEN_HOSTS`. Everything she plays also goes through YouTube.

2. Sites her code contacts directly to look things up (`FETCH_HOSTS`, exact names only)

- api.open-meteo.com and geocoding-api.open-meteo.com for weather
- en.wikipedia.org for encyclopedia summaries
- www.youtube.com to find the top video when you ask her to play something
- news.google.com for news
- www.bing.com for search and news

3. Pages she reads from search results. When she looks something up, she reads the text of the top 2 results (`READ_PAGES`). Those pages can be on any public website, not only the lists above. There's no site allow-list for that step.

- She only reads, through https and public addresses. It never downloads files, saves anything or runs anything, and web text can't write her tags.
- The code skips these sites: x.com, twitter.com, facebook.com, instagram.com, tiktok.com, linkedin.com, youtube.com,youtu.be and pinterest.com.
- If you'd rather she read only a short list of trusted sites, or none, set `READ_PAGES = 0`. She'll then answer from headlines and snippets.

******To set the weathers home location set in talk.py******

```
HOME_LOCATION = "Location Here"           # default city for "what's the weather?",  
e.g. "Seattle, Washington"
```

- **Warning:**

!!!!!! If you don't know what you are doing you can break Liz !!!!!

Make sure you have a backup copy of the file you are changing before you change the settings. DON'T use TAB | DON'T add extra spaces or delete extra spaces. Only type and change the settings and you will be fine.

These are the settings that control timeouts, retries, error checking and debugging. Each one is in the settings block at the top of the file named, so you can change it there.

Talking to the model (**talk.py**)

Setting	Value	What it does
LLM_TIMEOUT	(5, 120)	Seconds to connect, then seconds to wait for a reply, before giving up
LLM_MAX_TOKENS	400	Longest reply she can generate
LLM_KEEP_LOADED	86400	Asks Bionic to keep the model loaded for a day after her last request. Set to None to stop asking
WARM_UP_MODEL	True	Loads the model at startup with one tiny request (doesn't work)
LLM_TOP_P, LLM_TOP_K	0.95, 64	Sampling values. If the server rejects them she retries without them

Web (**talk.py** and **web_tools.py**)

Setting	Value	What it does
TIMEOUT (web_tools)	8	Seconds allowed per web request
MAX_BYTES (web_tools)	1,500,000	A bigger web response is cut off and thrown away
MAX_RESULT_CHARS (web_tools)	2800	Most search text ever given to the model
READ_PAGES	2	How many top search results she reads in full (0 = headlines only)
PLAY_FIRST_RESULT	True	“Play X” starts the top YouTube result
WEB_ENABLED	True	False turns off every web feature

Checking that a page opened (**talk.py**)

Setting	Vale	What it does
VERIFY_ACT IONS	True	Checks the page appeared and retries once if not
VERIFY_TIM EOUT	3.0	Longest wait for the page. A page that works is accepted right away
VISION_ENA BLED	True	Screen and camera looking on or off
VISION_MAX _SIDE	1280	Screenshot size, in pixels on the long side. Smaller is faster
CAMERA_IND EX	0	Which webcam “look at this” uses

Listening and timing (**talk.py**)

Setting	Value	What it does
SILENCE_DURATION	1.8	Seconds of quiet that end what you said. Raise it if she cuts you off
MAX_RECORD_SECONDS	30	Longest single recording
MIN_SPEECH_CHUNKS	2	How long a loud sound must last to count as speech
AUTO_THRESHOLD, THRESHOLD_MIN, THRESHOLD_MAX, THRESHOLD_MULT	on, 0.003, 0.05, 3.0	How the speech level is set from the room’s background noise

PROACTIVE_MIN_SECONDS, PROACTIVE_MAX_SECONDS	15, 60	Quiet time before she may speak up
PROACTIVE_BACKOFF, PROACTIVE_MAX_UNANSWERED	1.2, 10	How her waiting grows when you don't answer, and when she stops
AUTO_GESTURE_COOLDOWN	6.0	Seconds between automatic gestures
OPINION_CHANGE_COOLDOWN	600	Seconds before she can change her mind about the same topic again

Debug printing

Setting	Value	What it prints
LISTEN_DEBUG	True	The background and speech levels it's using
DEBUG_GESTURES	False	A line each time a gesture is triggered
DEBUG_ANIM (avatar.py)	False	Idle and automatic gestures as they play

Memory and backups (**talk.py**)

- **BACKUP_KEEP = 7** is how many days of daily memory backups are kept.