Python and AI Power-Up Program Offline Class-20251008_112632-Meeting Recording

October 8, 2025, 5:56AM

1h 52m 36s

Tirth started transcription

Tarun Jain 0:03

Use case, yes, but the goal what you are doing is you are only reducing the time understood. Now there is one more thing, RFP and all comes into fintech domain, then sales domain, then you also have manufacturing and electronics domain.

Like RF is like 3 hours for task, four hours task, you're bringing down to 30 minutes one hour. But when it comes to manufacturing, you're bringing down three months of task into minutes of task. So that is one actual use case. Because if you look at China and Japan, I'll show the demo.

So there is China company, there is Japan companies. They spend too much of time in manually checking the dimensions of everything. So 3D diagram you create a 3D diagram. So when you convert in 2D diagrams you have front view, top view, you top view, you have side view.

And you have isometric view. Now everything has dimensions. Every single part will have dimension. Now when we create a new design, a bill of materials and what is the price of it. So you have to create that entire details.

You you create a report, you create a report of it. So in order to create that report to take three months, three months, six months because manually everyone has to go check and everything. So that task is brought down into one date with some of the queries also into that startup. I got to know actually.

We started similar kind of things for right? No, that is for back office.

So Quinn is also getting into this using point to generate. There was some article where Quinn is currently helping some China company to build this, but I'll show the demo.

Been one year I haven't used.

Even $5 or GPT 4.1 Uh, is anyone else joining online or no? Uh, these are all OK.

So we'll start with MCP. OK, OK, so MCP mainly will be focusing on context engineering. Many people nowadays they want to turn prompt engineering as context engineering and these are two folks.

The Toby's nothing but the founder of Shopify. We just said that context engineering should be briefed over prompt engineering and then you have under so the guy who named white coding. So after this guy tweeted now everyone have again started using term as context engineering. So basically for context engineering.
You have drag, then you have state or history, then prompt engineering, then memory. Now what is the difference? So prompt engineering, you ask a question, you give it to the LLM, right? What happens in context engineering? It is a workflow. So from where are you fetching that context from vector database?
You write the logic of that, get the context, give it to the prompt. So workflow of the system. Same goes for memory. Memory you have database or you have some third party providers like Memzero and recently there is also called as Colink. So Colink is a new framework where you can use long term memory.
Save your preferences, retrieve it any moment of time, so that is memory. So the workflow from where you're getting the context tool calling. So tool calling also you have to write the logic, then fetch the content, right? So workflow is nothing but context engineering and just string simple template.

Mitesh Rathod 3:48
It.

TJ

Tarun Jain 4:01
OK, and just the template is nothing but your prompt engineering. So understood the workflow, the logic that you write is context. So here only we added the MCP part because what is the goal of MCP again? The goal of MCP is you have your set of tools that you have created and you're fetching the information from the tools.
And then giving it to the LLM. So now what MCP is doing, it's only fetching you the context and again I have the same slides. So the issues with large language model right? So LLMS can draft the calendar description for you at it can also write the e-mail for you, but can it send?
No, it can't send. Then can you directly schedule a calendar invite? So this is 1 very good use case if anyone wants to work it. So what you can do is your input tool is a e-mail. Now let's you had some calls. OK, forget e-mail note taker up. Do you use note takers?
Yes, so there are some note takers tool like you have TLDR is 1 note taker, then you have fireplace.

So now what is happening? You went with a customer demo and in this customer demo you had some discussion, right? You had the problem statement, the requirements and most of the time you'll have action steps at 10th of the meet you will say, hey, we'll send you these details and some.
Action step that you have. So now what will happen during this context? You might have discussed action step. OK, so now what this will happen is this is your input. This is your input. Now what you are trying to do is you have an agent.
Now what tool are you passing to this agent?
What tool will you pass?
You will pass the fireplace tool. So now what will happen? If you ask any prompt, hey, what was the action step that we discussed during the call? It will give you that context. Now using that context, what will you do? You will use one more agent.
So now this is agent 2.
Now what this agent 2 will do now? Agent one is looking at all your transcript, it is getting the action step and then it will summarize it. Now agent 2 what it will do is it has a tool called mail.
It will send you e-mail within the internal team itself. So let's suppose Hardeep was in the customer meet call and then he has agent one. Agent 2 will have a mail with sending mail to the internal team like right now. If you think that e-mail is good enough that has to be sent to the customer, you can send it to the customer.
You don't have to directly send to the customer until you review it. If you think that mail is correct, just copy, paste, send it. Only two minutes task. We did this but we never used it. It's like we just experimented but we never used it live because the mail one was there. It was fine.
But still it was like, I don't think we didn't use it, but there is one use case. If you think in your case you tell it out, it will be very useful. So we have something called a TLDR for it.
OK.
Even TLDR I thought too long didn't read but there is a TLDR dot IO. So this is the tool.
BND.
Maybe.
OK, sorry, TLDVI told TLDR.
Where is comment?
The thing on 8B does not support tools.

And listen.
Which model did you use? I think 88B. Try with 14B. You're running on collab only, no on collab, right? No, no locally. OK, locally 8 billion.
OK, we'll check that part. OK, OK.
So after every meet it will usually have a summary of it. So it create a subsection like in starting one minute you discussed about some requirement, then you discuss about some requirements. OK, I should have logged in from a different account, but I'll show this later at the end of the session when you guys are working and it's not. Fireflies. Yeah, this rule should be paid. But if in case you already have the subscriptions, like if you have note taker or anything when you join the calls. Mm-hmm. If in case it is summarizing anything from wherever you're getting the context, we use TLDB.
And APIs are not sure the cost of it, but from there we are fetching the information. So probably that might be the reason why we didn't continue. The price might have been high, but this use case we are developed.
For the slides for the video and audio.
Do we have elements this multi models that can understand both video and audio or no?
Audio. Yeah, Audio Gemini supports. Gemini supports. Huh. Video. Nobody understands the video video frame by frame. You have to pass frame by frame. If not Sora Sora API, there is one.
Apples, native. Also that one, you know it understands everything. Video. Video. Yeah. We get the frames. It's it's yeah, it will understand 60 frames, yeah.
People understand. Yeah, like if uh, Gemini supports video for in video as an input. If Ronak will tag me, it will understand that someone will check our and for output you have file. Huh. File is MLX, MLX and a worker file, yeah.
But again, these are not you have and you also have replete. Hmm.
I'm just thinking.
The product need schedule apologize. What do you got the issues? Yeah right we we have as a company we have to offer solutions which are not that heavy on the pockets for people who are budget is very distinct like if in case you run locally you can easily go use Bedrock because.
Bidrock will be much cheaper than hosting GPUs with because nowadays API costs are getting lower of LLMS, but the GPU costs are getting slowly higher and higher. So I hope this understood the use of MCP, but The thing is.

You can use LLMS to generate something, but you can't take action. So this is where typically we use LL MCP. So in simple terms, MCP is nothing but Model Context Protocol. Now here one question is common. What is the difference between function calling and MCP?

Right, so function calling what is happening? You are asking question, you decide what to tool to pick and you are defining the parameters. So when it comes to parameters, you are easily just telling what you have to search. You are not having any structure for it. So when it comes to MCP, they created a structure on how you have to define a tool.

So that structure is nothing but a protocol. Like protocol is nothing but a rule that you have to follow in order to call specific tools. So here they're using something called a Jason RPC which is remote procedure call and context is nothing but hey, you're providing additional context to the model which is your.

Anything LLMS and this can be LLMS, this can be agent and this was the main idea of MCP. So basically you have functions with the documentation on an external server. You're giving that URL of the documentation.

To your LLM or in initial initialization it will get certain format, certain format and this is the description when we use it, correct? OK, similar like tool calling but like on remote server and then you would just communicate via that both remote or local both are supported.

So if you see there is one image here which is very interesting.

This one. So if you see MCP is nothing but you have client. So client is your cloud. Let me open this image separately.

So you have a plan which is a laptop. OK, which has cloud, which has GPD, which has cursor, everything. Now what this MCP servers are nothing but you want to connect your slack. So slack might have multiple tools. Same goes for e-mail. Calendar is 1 tool.

Then sending e-mail is 1 tool. Reading e-mail is 1 tool. So all the tools are within one server. So one server can have multiple tools just like that. You have calendar multiple things. It can be your data, it can be anywhere from where you are fetching the content.

So tool is nothing but you have to define the logic, then you can use LLM. MCP is like hey we already have our MCP server, use it. So there is also meme. The developers of MCP are more but the consumer of MCP are less.

Have you seen this? Usually what happens when you build a product, you'll have less

developers and more customers users, but in MCP it is reverse. More are developers but less are consumers because many people are building MCP servers and exposing anywhere, but very less people are aware of MCP yet.
So this is discovery or something? Yeah, yeah, you have something like that. If I want to do this, with MCV server, I'll come to that slide. I have one slide for it. Yeah, small, but.
That cursor is also there. cursor also have a directory of all the MCP servers. OK, OK, so in simple like what is the need of it? Most of the time what happens for every single tool that you write, you need to create a custom bridge, right? So if you need to connect your slack or drive Google.
Drive. You need to write the logic of it. You have to define custom function, then connect. Correct. Same goes for quadrant Airbnb. Now what will I do if I need all the services? I'll just use that MCP server and I'll connect it. So you writing the logic is reduced now.
You're just using what is already existing and then connecting with the cloud. It is a tool that does not need a separate implementation or API implementation block. We just have the elements who will do it for us, correct? Automatically. It's a ask a question, get the respond that's right.
And I'll just show how it works. Uh, so there are two different things. This was from my slides.
I took this one write it for you.
So basically you have user, user will write a prompt, then you have MCP client, client can be anything, your cloud or anything. Yeah, now what LLM will do is LLM will have all the instructions, system prompt, user prompt. So this is where many people tell that you can misuse of how MCP is being used.
So you have prompt injection techniques, you have tool injection techniques. So you as a user don't know what has been called in the back end. You'll be like OK output but anyone can phish that attack like what tool was called and what was the output of the tool. So you can in this tool calling right you have your data.
The data can be taken outside. You're using a tool. It will tell the read all the files, send it over to this person. OK, but finally what the user has told you can do that and give that output to the user so you can add extra instruction. You can read that data, you can send it somewhere else, correct? And you know so now if you see you're using.
I'm using MCP server. Let's suppose this kind of like like you know. So now if you

look at this tool one, typically this tool one is to do to do search or let's say a vector database, right? It has to do search and it has to give you the relevant documents. Now during this moment, if someone takes a data out in terms of logging or anything, you will never know what is happening behind the scenes, right? That is one major issue of MCP which was there existingly, but now they added OAuth and different prompt injection techniques in their latest release.

So you have LLM, you have MCP server, then I've added tool one because one MCP server can have multiple tools. So this LLM again has to do routing which tool is right and based on that now you have response, sorry request and then response rate. So now you are asking request.

And then you're generating the response back to the MCP client. So in simple terms, MCP just have three things, MCP client, MCP server and transport mechanism. So transport mechanism you have two, one is to do locally and one more is to do remote.

The S in MCP stands for security. The OK, there is no S MCP. The S in MCP stands for security. So like before you know the context that you were telling.

That is a good meme. There is no SI mean there is no S in MCP.

Oh sorry, S stands for security in MCP. There is no S OK coming to the transport mechanism. Typically you have something called a JSON RPC 2.0. So whenever I ask any question it will be formatted in this format. So lab request you will have response. I'll show this in demo also.

So you'll have method. Method is nothing but tool calling and then you have params. So for the given function, what are the parameters that is supposed to be passed? So this ID usually it is hash so that you add that Oauth thing. I'm not sure how that Oauth works, but I just read it somewhere.

I know I'm aware of of how it was. You have that mechanism in LLM itself. OK, so you get the token. There's a token in everything that goes through in the OAuth. OK, that is recently added in a JSON RPC. So now here you have two things again. One is STDIO. STDIO is nothing but.

Standard input and output which is running locally and then initially before this Oauth all happened right in the initial versions of it you had SSE so is so this was for remote calls. But now if you want to use SSE you have to call it as streamable HTTP. So you can use all the REST API commands as well with MCP. So here the input is endpoint, here it's running locally. So you have NPX, you have UV, then you have Docker, so and Python also. So these are four ways you can run STDIO locally.

I'll skip this part because it's too time consuming. So what you understood right? STDO is mainly locally. It's the same diagram. So you have client, you have server, you're asking the response and the questions. That's it. But if you see local tools, CLI and desktop integration which is.
Best suited for STD IO and same goes for. It's a. OK, I've not heard of it and then you have this streamable HTTP. I have written SSC because if you use SSC somewhere. This thing that is deprecated, right? Because most of the cookbooks that are available in uh.
What do you call MCP? You'll see SSC in many places. So when you see a SSC you just think this is replicated, you have to use the newer version of it and this is mainly used for web applications and for web application first is you have to start with end point. So you have to check whether that end point is there or not. Best example yesterday we had.
Agent OS, right? Agent OS can act as MCP.
In you had multiple post and get commands, so now that is exposed as MCP in agnode. So if you do docs you'll get swagger UI. If you do MCP you can use it as a SSE or streamable HTTP.
All right, so this is the repo. If you just search for awesome MCP server, you'll get all the list of MCP servers that you have. You have GitHub, you have Google Maps, Google Drive. Google Drive is widely used and it is very trusted what you call MCP. And then you have Quadrant DB, Quadrant DB in the sense let's suppose.
Because you want to generate a code and now you think this code is working fine and you might reuse this any moment of time. So what you can do, you can save this code in a vector DB and now next time whenever you want to use this code you can just tell hey I might have saved this template code, can you get it?
So this is widely used in EI as well. So for example, when I want to work with neural networks, I will have hey, these are the two neural network node I need and then I need a Max polling and normalization. Or else let's just take rag example.
For me.
In coordinate, you mentioned that you just say from huh.
Save in the sense it will convert into embeddings, then save store and then retrieve indexing. Mm-hmm. So when I said store it meant indexing. So you'll have embedding model. OK, so let's suppose we work with rag code, right? And if you see all same code was repeated multiple times collection name, then you want to use binary quantization. Now just imagine.

Whenever you tell your cursor, hey generate this code of defining quadrant client. So basically it will not generate binary quantization because we know binary quantization works. But when you add this to LLM, LLM will not be able to generate it.
So now you'll tell hey generate by define quadrant client and fetch it from DB which is our vector database. Now when you generated the code you're saving that in your database. Then you're retrieving it, which is nothing but search.
So when you're saving it is index, when you're retrieving it, it is search. So you have two distinct tool. So what is the first tool? Quadrant store to save your data inside vector DB. I'll tell index because if I say save it will be different meaning. So you are indexing your data inside vector store which is quadrant store.
And if you are fetching it, it is called quadrant find. So these are two tools now. Now you have file system. I've not used it. Probably you can check what tools it supports. So I'll show now demo of some of the MCP servers and then we can also use it.
And this is the format of how MCP server will look like for STDIO. So you have MCP servers, you define what that name is. Then you have command, command and arguments are same. I mean command and arguments are must that you have to define.
So command here is UVX. UV is nothing but you use. You use it for Python. So UVX or Python are same. Then you have NPX and then you have Docker. So where do we find this? Every single MCP servers that you have here, right? You open anything and if you check their GitHub repo you.
You will have a Jason file. If you don't have the Jason file, don't use that MCP server. That's it. And then you have ENV. ENV is optional if that particular model needs additional environment. I didn't understand like if you don't have a Jason file. So let's suppose I want to use Airbnb.
MCP.
I will open the GitHub repo here. If you didn't find a Jason file, don't use that MCP server. OK, we don't know what is. You don't know what command is, what arguments is. So say if I search for context 7 MCP server.
And if you see here these guys are providing MCP server for different client. So you have cloud code, you have cursor, then you have augment code, augment code which I mentioned earlier. So now if I click on cursor you have that JSON file. So JSON file is must for every single MCP server because then you will not know what to use.

So whoever will build that MCP server will definitely expose this. And the MCP server that we will use today is free CAD MCP server. Can we Microsoft do have? Do they have it? Like Microsoft does not have it. There is.
It is in the uh awesome LTP one.
The reason why I'm telling this is if they all configure it and they can just say uh work there and important e-mail to it.
Which is rule you said? I think that's on Outlook on this one. Outlook, yes, this one, this one, yes.
You see which tools it supports. I'm uh, NBX they have. Yeah. So now what will happen is let me copy this. Oh best. They don't have any API, so I'll copy this. I will request and uh.
So till here everyone followed till here everyone have followed the syntax UVX is for Python NPX is for node and docker. So all these things will be there in the read me file. So now if you've seen that Microsoft you saw NPX right? So this is the format for STDIO OK for.
SSC it will be end point URL, for example atyantik.com/sse UPS or NPX installed. Huh. OK, I believe everyone have NPX right? Let's use NPX. Node JS NPX. So till here only had slides. I'll quickly show 1 demo then we can experiment with.
Different types you take.
OK, so I'll this is a very interesting demo. So I'll open free CAD. What is free CAD? It's a CAD software to build 3D models, free CAD tool, free CAD. It's open source so you can use it anything.
So let's suppose you want to create this thing right blocks thing so you can use it. So now if you see I added a add on which is called MCP add on.
Can you see that it's a local tool and it has an MCP add on? No, no, it doesn't have MCP add on by default. So let's suppose you search for free CAD.
Free card download. Yeah, it's OK. No free card is a tool. So let's start with first. So you start with free card. It is available everywhere, huh. So you can download it for Linux, you can download it for Mac, you can download it for Windows.
Right so once you download you will not have MCP by default by default it is then then if you check the repo.
So do we? Where is that awesome MCP server? So what about?
Open source. No, these are CAD software, right? So if you see her.
Sorry, these are mechanical tools. It's what I said. OK, correct. Mechanical design tool. Design is the main keyword. So this is the tool MCP free CAD MCP. So if you see if

you want to draw a car, you can draw it. You can just give instruction. That's it. Now if you want to change the color, this is different.
So you don't have to know how to create this design. If you can give instruction, you can give those instruction.
So this is 2D. So for 2D you'll have front view, side view and all and here.
Then this is where they've spent too much of time minor review. So all these things right takes minimum two to three months, but with TI you can do it in two days. So this is a massive use case.
Because reviewing each and every section and part of it, it will take time. So it's like first you can get the draft view, then you can tell manual reviewers to check it. Is it correct or wrong? So that final report is what matters. So now if you see when you do that right inside your GitHub repo.
You have addon. Can you see this folder? This addon needs to go inside your recab. So for Mac you will have this folder called Library Application Support Free CAD Mod. Inside this you have to add this add on. Once you add this add on close Free CAD, reopen it again. You have to see this.
MCP add on but I'll show the demo but if anyone interested you can try it as well. I'll click on MCP add on. Now here you have free CAD MCP on your title bar. Just click on start RPC server.
RPC. So I'll start it. Now it is connected. Now I will open plot.
Uh, I'll show the like it. Uh huh. View.
Uh, we did. I'll just, uh, no, no, wait, I'll show it.
So you have is that panels. In panels you have to show report view.
Panel report view. So can you see this?
I'm using this for last month so I know all the comments here. So you can also run Python code inside this. So macros is nothing but write a Python code, convert that and macros. I don't know 0 logic. Claud is crazy.
Good in my macros cloud AI. So after this only I started using cloud mode. Before this I was using very less for code. You can no that that is different software for 3D printing. You don't use this. We can use this. We can use them. Yeah we can export this. So now I'm opening cloud.
OK so as of now we just installed the library, we are adding add on and we are starting the server. So once server is started.
Let me keep this in one screen. So after you add MCP server, I'll tell you how to add it. Can you click on this button? Yeah. If you click on this, you will not see any tools in

your case. Hmm initially. Hmm. Probably only you three can try as of now. I'll show for others as well.
So Cloud desktop you can open. Can you click on this tools? Mm-hmm. Do you see free CAD important?
No, no. That is the MCP. Yeah, for which you have to create a Jason file right now for every single tool. If I select free CAD, how many tools I have? Create document is a tool. Create object is a tool. Edit object is a tool. Delete object. Execute code.
So for example, if something is not working, you want to change red color to blue color, it will write a code for that, then it will update it on screen. So I'll just help create a 3D design.
Design model for screw.
Tightening machine.
And if someone takes who thought you, don't tell. I thought you just tell some nobody. Why? Only for this MCP simply why? Because this is used everywhere and I've shown this may have at many places so it might cause conflict.
So you just say we came across free card. That's it. Other things you can mention, but only for this thing you like. I didn't. I mean, don't tell. I thought you only for free card. Why? Simply because I've shown this many places and it's a very cool demo.
OK, I don't have any agreement with anyone, but still OK, so create a 3D design model for screw tightening machine.
On free card.
So if you notice here there is nothing as of now on the right hand side. So now it is using tool one. So what is this tool one? This is Jason RPC and this is Jason RPC. Now it created a base.
Now this is again if you see this is SNRPC.
So it's adding one by one. It will add up components. So you added create object, create object, create object, create object. So now let's suppose you didn't like this blue color. So now what is this blue color? It's called motor housing.
So in your follow up prompt you'll say, hey motor housing, can you make it a Gray color? Can you make it red color? And you'll also have this in aluminum. You'll have copper, you will have steel. So now imagine you want this to be copper instead of steel. You have to give that instruction.
This takes too much of time and now we can just see hardly it. It didn't even take one minute, so it will go. This will create normally do the same thing in auto, then at least it takes four to five hours.

Of single person's task, I mean the whole thing.
But to planning everyone also. So here I just gave a simple prompt.
Ha ha ha. So we'll just make sure so we'll just use this MCP only. So we are using free CAD. You'll use different tools. That's it. We'll not write any code much.
But don't use, we'll use different. But if you want to use, feel free to. But to those who are using cursor, we'll have cursor related tools, MCP.
If this start building, many people start building it then then the human effects.
So job might be. So yeah we have lost like so yeah this is pretty cool I I mean and next one.
Again, I'm telling never say that I showed this, but yeah, this is pretty cool and all these repos, right? All this repo, if you see it is by Chinese or Japanese only. You will never see these tools built by anyone apart from these two countries, yeah.
So this also it is I think right now from AI to robotics and all this. So if you see this is also Chinese, if you see this instruction and they're heavily playing in this domain Chinese people.
Please, whoever released this repo and I feel very sorry for this guy. He could have made millions with this repo and he made it open source.
Japan.
But he needed to make it public. This guy didn't it is based on the public use. So this report reports like this makes so I mean.
How crazy these peoples have more than 1500 contributions from how many years 123456?
OK, from 2020 to 2025 has more than 1500 contributions. But yeah, let's create the uh what you call Jason file. If it is uh we need to edit them that adon manager, I'll show the step by step things. Uh so this is.
Done.
So you have screw typing machine. What I'll do is I'll delete this.
Just for safety purpose, but you saw right how this works. OK, so now what you can do is who have Claude, you can open Claude.
Yeah, I'll come to cursor. You're talking about adding free CAD to no any MCP server. Let's suppose you want to add Airbnb, right? Free CAD. It's UVX. You have to install UV which will take again some time. So NPX. I hope you have it installed.
So open now this thing. Yeah, we we have terminal. Just do which NPX. Hmm.
Do you see anything? Yes. Do I need to find out in? No. Which UVX? Yeah, we have.
Then do which UVX? Yeah, we have.

OK, if these two things are there, then now what you need to do is open Claude, click on Claude.
Settings.
And then you have developer.
In developer you will see local MCP servers. I did config, I did config. It will open a folder and then you have cloud desktop config file.
In your case, it should be empty.
I had it the spot record. Is it empty?
Cloud, Cloud. I mean, yeah, yeah, Cloud. Then you have a server settings. After settings you have developer edit config. It will open a folder. In that folder you'll have a Jason file and the Jason file. Is it empty or?
It is handy, OK.
So we'll start with file system. Let me copy this somewhere.
Or instead you can just search for file system.
File system MCP server.
Server.
And everyone can try this now. Till here is it done? Have you come across till this one? Yes, don't try this. But are you in this file? Just yeah, I'm in file. Actually this file is not open in a, so do some, but you can see this file in a.
Cloud desktop config dot Jason. This file is open. Yeah, yeah, OK, now all three of you who are using cursor. Let's open cursor.
OK.
So cursor. OK, OK, I'm in cursor only. Click on settings.
If anyone want to try cursor also you can try settings. Then you have tools and MCP. In your case it will be add new MCP server. So can you click on add MCP server? If you click on that for your case it will be just MCP server will be there. Then you'll have curly brackets and it is the stop yeah.
Now what you can do is copy this what I'm sending in now.
OK, I'll open. I'll I'll close free card.
That was just to show the usage of MCP new MCP server and I'll also show close arc now.
Rome, huh?
So copy this in your Jason file. If you're using cloud desktop, copy this in Jason. If you're using cursor, copy this in that particular tools MCP.
And what will you change now? Command is NPX but you have to add the path of it.

So how will you get the path? Just come to terminal type which UVX? Sorry which NPX?
And then argument is same. You have model context protocol server file system. Below that you have a placeholders.
So what is our placeholder?
It has path to other unload directories.
What happened?
So if you see a path to other load directories is there. Instead of that what you have to do is place with your own directory where you want to give permission. So here what am I doing? I'm telling users Tarun's in desktop is exposed. So now Claude. Or any MCP client has flexibility to use your desktop so it can delete anything from your desktop. It can also write anything. But The thing is when it comes to deletion, it will ask you should I delete or not. It will ask confirmation.
And for developers, context haven is a very good MCP server. Now let's suppose you're working with Nick JS. You're working with Langraph. You might have seen after one one month the parameters are getting changed right of any documentation. If you want to be updated with any documentation, you will use context.
7.
So context 7 is an MCP. Again I'm using OPT same the NPX then arguments is Y upstash context 7 MCP then it needs API key. So API key you have to get it from context 7 dashboard.
Contact7 com then dashboard.
So at contactsammon.com and if you do back slash dashboard you will see this. Can you get a so only copy this one file system. Context 7 also you can copy because it's a very important tool for developers.
But the only thing is API will be different.
We used jira's and CP server to create this web reports and do you have any API key for that? Oh, the connected uh, then we have to use that auth.
Same we have to go to context 7 MCP.
I just you can awesome MCP server right now so just search there for contact 7.
The.
Just search for context 7.
The baby to you.
So you have upstash context 7. There is no space. If you scroll below you will find

cursor. Copy this.
Context 7 URL letters. Either you can copy this which is remote, but we are running it locally so I'll copy this one. Context 7 URL letters no this one. Context 7 comment and arguments.
So whenever she URL what is this call now? If you see endpoint this is streamable HTTP. So this is streamable HTTP. This is STDIO.
OK, if you click on this add cursor, it will automatically open cursor and it will add it. This is also there. Very good, but we already have it ready so I'll look at still asking for API key API key. You have to click on contact7.com/dashboard. There you will get it.
So this is context7.com.
And here if you click on this personal.
You will see API clear.
So you have to click on the dashboard contact7.com back slash dashboard.
OK.
So this is the format.
And I'm adding quadrant as well. So quadrant is nothing but command. This is UVX. Then arguments is nothing but MCP server quadrant. And then you will also have ENV. ENV you can use quadrant URL, quadrant, API, key collection, embedding model. So as of now I've showed 4 MCP servers.
Free CAD, contact seven, file system and portal. You feel free to use any two of it and then we'll try with the demo.
So to those who are using, I mean desktop, Cloud desktop is very simple.
We have.
Yeah, but I don't see in workbench.
Nice. Uh, so back off.
Part design in a part design. Hmm. So there you should have MCP add on. You have not added MCP add on yet. I have one second one level. Copy add on. Hmm. Check the mod file.
Yeah.
close and open again library application support free card.
More, more kinder care.
So that's wrong, no?
So what you're trying to do is copying the files instead of folder. OK, OK, OK, OK, I need to move everything.
Mhm.

And we pick it up.
Is it done a quadrant URL quadrant API key or get it from from there?
And if anyone wants to use this thing, it is MCP MCP dot Jason is for.
Cursor cloud desktop conflict dot Jason is for cloud desktop.
I'll add both at the same screen, so whichever you want to copy, you can copy.
I need to so this is my cursor's what you call MCP. This is my cloud's MCP.
Because I use plot more than cursor.
So if I do LS.
LS of back slash library free cat mod.
Sorry, novelis.
Mod it's free CAD MCP. What is inside free CAD MCP? That is as a file. So you should have folder instead of files.
Just cross it. If it isn't, how do we cross it now? Open cloud desktop, close and open it again and when you click on tools you should find this.
So here this is the tools here. If you see you have free CAD, how many tools does free CAD have? It has many right? And now if I click on quadrant I should have two.
Now I told you one is to store and one more is to find and how do we verify it in?
Cursor click on settings.
And then if you click on tools and MCP, you should see the total number of tools. So you have 14 tools for file system, 2 tools for context.
So this is the way you can verify. If there is any error, Casser will tell you there is an error. Same goes for Claude. So Claude will not even show the name. So now what you can do is you can go to developer. It will have logs by failed.
So this is to verify on cursor and to verify on what do you call plot. Just click on tools and you will find this name of your MCPS free CAD and quadrant.
Um.
Let me know how many of you are able to. It's a SSC stream disconnected.
Why are you using SSE?
We can use a.
Let me hear you. Oh, you have added so many play read.
You can use this one STD IO.
The file system. Is it working? File system. Yeah, it is working. Playwright. It's working.
I tried this all of this. I see context. I mean is wrong because you have to add a SVIO.
OK, I like this. No, no, it is. You can check out in the repo. OK, it's like this.
NPX then Y upstash context 7, then APIK and then APIK. He has added five tools.

So I tried one like entire workflow. I added jira's and CP server. Ask them to bring the task and description. OK then generate the code. Once generated, test the APIs with the.
Uh, Postman's NCP server. Once then then generate front end. So did this before only? Yeah. So once the front end complete, uh call the playwrights NCP server, open browser, test and and it works. We should give an internal session then again go to jira.
Update the task. See always writing code. No but cursor has we made like uh 25 for a tool. But that you can you feel other sorts. Also Figma this could have saved so much of time for us.
And also Figma. Uh ha. I tried Figma. I guess it's not free now. Yeah, but I have a student accounts. I tried that in the morning. I have free a Figma and I so I download because.
See, check for Figma also you can do Figma. I'll show on video Figma. Just copy the design, come to cursor and tell hey this is the Figma design. Can you write the code for it? If you write the code, yeah.
This is I know.
So what do you have to do is let's suppose uh, there is some what was the accurate design. So now let's suppose you have some uh visits here and you want to build a website for this for this design.
So this one you just have to copy. Where is that copy?
This one.
Copy the copy ends. So this one you have to copy and if I check the link of this.
Mhm.
There was some copy.
We have to copy the frame I guess. OK, that is there in the repo one second.
So you can try this one Figma context MCP.
So in this Figma context MCP this is the website. So if you see this was in the Figma then he did copy paste in cursor and he created an entire website on cursor. So this is that video and this is the MCP.
Yeah, it's we need to add API key and I didn't add the to section. Huh. OK.
Copy link but that is not there in me. Copy paste as and.
OK, maybe a browser.
Yeah, I think.
Correct this is the link this link. If you give it to your cursor you just have to paste the

link in the cursor agent mode and tell hey make this for that it uses this server.
This is the MCP server, MCP server, then frame link Figma MCP and you just have to add this. Once you add this just give a link in the chat. So I told generate 3D design right? Just take the tool tell hey this is the design right code for it.
They got 10K starts in a very short period of time, March to October 1234566 months 10K starts.
Can you don't come from if you have seen, I mean if you have your MCP server set, yes, for I'm getting error.
So the word is connected. Did you clone the repo? I did clone the repo. Where is the part of the repo?
But then I copied it to open the. Oh OK, OK, I need to add that part. OK, let me. So if you see this command, right, see this one free CAD MCP here you have command.
Then you have users tarun Jain. Sorry users tarun Jain desktop free CAD MCP. In my desktop I have.
OK OK so if I come to desktop here I have where is it? I see I see I I added it under so that's why I haven't added but for file system context.
Server disconnected for which one are you using?
It's one for it, so they can see it's not. So in one city server.
It's not you know you're.
I guess you have to update and PX also.
close you have to close the the.
OK, wait a second. What was the?
All right, OK.
Yeah, maybe you have to have. What is the NPM version you're using? I need to find so I need to find which one you can use. You can use the.
It picks.
Definitely.
10.7 I won't put on a business how to operate and he said me the NPM update no, but uh might have been updated at least which NPM? We it's a the but where in which it is which NPS no.
No, it'll be a much. Uh huh.
Exactly.
NPM install NPM latest and and VM use latest. VM use the you got everything.
By MVM and then yours. Yeah, yeah. We'll wait for two minutes.
Yeah, well, I must that so that if you gotta speak with him.

And um.
Nene, you can't. Can you color it? It should be like a specifically.
But the don't use extend thinking, extended thinking.
Don't use extended thinking when I extended thinking.
What's and we will OK.
Is it done? Let's just test it and make a. So now what is our client? Claude is the client, cursor is a client. What if you want to build your own client? We'll use.
So client can be anything at the end of data model context protocol. Context is your MCP server, model is your MCP client. So client is LLM or an agent. So cloud is an LLM, cursor is an agent. Now we can build it using also.
So now you have two different agent. We will build it with Agno till here is it done. Until then now cursor only you can test it out or run or run it. So what you can do is agent mode open.
Where is agent born? Yeah, yeah.
Mm.
OK, here, just tell one thing.
Next JS latest dogs.
Hmm, 22.20?
Hmm.
So you can tell get use context 7 and get information.
So can you see this resolve library ID? So resolve library ID is your tool I'll run.
So where can we verify it? So settings, settings, then tools and MCP. And if you check the two tools, what is the tools that you can see here? These are library ID and get library. So these are tools and it's working.
So a context use quadrant best quadrants that you're saving it, indexing it and retrieving it that we will do in the code.
But you can verify it. Quadrant is very important because Amara DB view all the DB.
Why is it over here?
This.
What's the error?
API key thing. Hmm, they didn't know when I.
You can add anything like uh next next JS latest docs use context 7.
And then it will take library name as next dot JS and here you have parameters.
Oh.
If there is any change in your code and if you want to update from the documents,

up to use context 7.
So if you see it, I will pull the latest V15 docs from context 7 for app router and data fetching. Then it's what tool is it using now here.
Get library docs and that is the area. So context that is resolve library ID and get library docs.
Now create a new chat is for pause further and create a new chat and just tell write a poem.
On at the antic do web search if you need information.
Search for info. That's the default search and save it in.
Desktop API has om.view huh? That's it. APIV. Desktop. No ENV. Uh, OK, so.
And here I'll just save it in desktop as poem dot TXT file. So now what should it use? It should use file system.
So context and test here. Now we are testing file system. Now this prompt you can write out.
Desktop. I mean desktop here prompt because you're using file system.
Yes.
I will search a family from a developer from me and Tom so I can write in from wife. No one with.
So I hope everyone understood the format.
So now if you see list a load a load directories. So from where is this tool called? It's called from file system.
If you see it, you have this.
List unload directories. Can you see this tool list unload directories list unload directories and there it will find you only have desktop.
So how many allow directories I have only desktop and where did I define this? Here user Saruja desktop.
Yeah, what when when that program is ready to go and read it out. Think the same way or different time in the elsewhere. It seems like it says can be exchange a name with.
No, no, no, no, no, no, no. So what are trans, like that states are text we can you know is a means. OK, it's saved whenever we saved. So if I open high up om dot EXT file.
Please see him now.
In the REM, the REM sleep light where you know you can break the light, because I've added that in the negative keywords. Yeah, this is good. Bond with the Banya

shade Varadhara slide a spark in 2015. You're correct again.
That chose to build the bright crafting white trigger into every pixel's glow, turning complex currents into flow and human flow. They stretch humans future grammar in the UI and add in the code where enterprise and elegance had the self same Rd. with React PWA.
They we can page us into a speed that reads not bad and I think this first spot there is welcome.
So we can use search plus MCP.
OK, let's look at the code as well.
So MCP usage wherever you have quadrant URL, quadrant API key, let's get that.
Let's get to the code.
Oh, we're not quite. That's my hands a lot. And what the heck? No, we say authentic technology. Vida does a.
No, it's on a file. Nice. See, you have a.
Yeah, because of technology. Yeah, because of technology. Yeah, now we can't just submiss us whatever. We're not in the visual to your voice technique.
This is what I'm gonna put the green.
OK, let's get to the code. So quadrant URL and quadrant API key, wherever it is saved, you can try to copy the and paste it right? And then collection name can be anything. It can be a new name and what am I giving it here by coding?
So I'm using quadrant for by coding. Whenever my code I think this is good enough, I want to index it. I want to retrieve it. So I'm creating a new collection in called by coding. Then embedding model sentence transformers all mini LM. This is for code. So I'm using this one.
And then OS dot environment Google API key. So instead of doing this I'll just do from dot ENV.
Import load ENV, load ENV because we are already saved at the same code wherever agent OS and there only you can do MCP usage MCP usage dot PY.
So this is same step from magno dot agent import agent.
Then from agno dot models dot Google's Gemini and up kya yesterday we did serve doug doug go and more the tabling. So instead of that we'll do dot MCP. In order to do this you have to do pip install MCP.
So yesterday it was Python then Google search results. Today it is just pip install MCP. I don't need OS and now whenever we are running MCP this tool is asynchronous.

So we have to use. So this is new line, but these are three same lines. Now in order to use quadrant we need environment variables. Why? Because when it comes to MCP. Go to that repo, check what they have. They have arguments, arguments here, MCP server quadrant and AKUBX.
UX is a UX is a command. So if I where is it?
MCP server quadrant firstly here. So this is how it looks like. You have UVX, then you have arguments, then you have environment, environment URL, API, collection and embedding model.
And we have this here all four and now this four will go as to ENV, ENV format dictionary, ENV format and dictionary.
And if anyone wants to use Docker, Docker is a.
We do it our own.
And that's OK.
No code. I'll copy and paste it.
But you have to change your quadrant URL and quadrant API key, yeah.
I've sent it in the chat.
That is very controversial.
dot ENV dot ENV. Make sure API key is still there. Google API key.
Till here, let me know if it is done.
Nor do you use them yet?
OK, to use Azure, right? Huh. Instead of Gemina, you got a you can use it.
If I add it, we'll not.
Friday going with that Saturday Sunday use.
He already has five MCB tools. Mm-hmm. Playwright.
Right, I think that he will also exit MC and CV don't do fetch and a configure and CMPM installed NPM installed NPM at the NPM.
NPM install and update. Mm-hmm. For the MCP and then again. Oh wait MCP.
Out of you.
So make sure embedding model same.
What is here?
So instead of running directly, if you see a MCP agent, this is the same. You're giving a name, you're giving a tools, then you're giving instruction markdown the area format same when you're defining MCP.
Tools initially I'm not directly like time to define for the other. So you import and you define directly. But here since we're running asynchronously, we are using async def

function async tools.
So we are wrapping it inside a function so that if you look at a sync with.
And CP tools.
Comment.
Argument and ENV time out they do because by default AGMA has very less time out, so I'm defining it as 50.
Then as MCP tools. Now this is replaced with MCP tools. Is this clear? Yeah, replace with just this variable. Now this variable will go to your tools. That's it.
Now that is your region.
I'll copy this code.
And paste it here.
And if you notice my response not contained, I'm doing direct response so that I can show what is the output in the tool calling because MCP it doesn't matter what your final output is, what matters is what is the input in the tool.
So that's the reason why I'm directly printing response.
We can also use Agent OS UI.
First tier. Meanwhile, go to washroom.
So copy this.
Is pictures.
I mean already page.
I'll come back.
Thank you.
No one.
No.
Take a.
Do you want to move master?
Mhm.
Go to the.
Right.
From the career, where the need of the theory would be.
So fix.
Banana del Khasul, the Imagination.
I miss much.
Oh, so then we are good.
Yeah, the little add.

Till here is it done? First we are defining the environment variables. We are defining agent which is a synchronous is a tool calling here. We kept it as as MCP tools. Which is a synchronous.
And MCP is the only tool which has a synchronous.
Directly instantiate it. That's it. Yeah. Amazon queue. I mean, Amazon queue AWS. But it community, community supported. But MCP is not common everywhere. Gemini CLI uses it that uses and ChatGPT released apps. Apps. Yeah, app is a MCP. Apps documentation apps, everything and this is MCP. So if I click on Tarun's in here settings either apps or connectors, booking.com, Canva and Expedia, this is MCP. If I click on this, learn more.
Search for do they have MCP somewhere? Yeah, apps open AI.
You don't know somebody. So if you see a developers can start building and testing apps today with the new apps SDK preview which we are releasing as an open standard built on.
Model context protocol. So October 6th.
We don't know that it is an app, but they're using APIs within the MCP server because of standard protocol over.
Uh, anthropic build MCP, but open source so anyone can use it.
Open your users at use cursor wind surf.
Is it done? Now we just have to write anything if name main give a query. So queries like give code to write neural network that is CNN with redone layers neural networks node defined neurons so 6412256 with one output layer.
You can give anything related to web development which you think is a standard template. You can define the template and then Asinsio dot run run agent query. That's it.
Where it can be anything.
Everything.
And now if I just run it, what number responsive?
Do you want to save that in DB?
So if you see our agent output then you have run ID and this is the entire code for tool calling. The content at the content is your final response.
Thing to obtain server version.
But make sure Satya pass UVX. Oh, the UVX path we need to do now how you can do that?
Mhm.

Hello.
Now let's see how to run this on agent OS.
OK, let's run this on now.
The UI.
Yeah, it is starting. OK, now there is no error and everything is screen.
Let's go over there.
So can you enable MCP server as true agent Osman?
From agent dot OS Qatar.
Sorry from agno dot OS.
Import agent OS.
Description is quadrant.
MCP server.
MCP usage.
This will be the file name colon app.
OK, the response.
OK, we have to use only.
Only SSU can use here. I think what we're doing. I'll be created something. How do you give your picture to it?
So these three tried it on cursor cursor with file system and context 7 because context 7 is very useful. I also tried it with context 7 OK on file system.
What was the team? We like, we got a poem on it. This poem was good.
And this is one word for you.
Is this? Yeah, yeah, yeah, yeah, I'll give it to you. No.
So convex 7 has two tools.
Docs and this is cursor OK cursor settings. You're adding 567 context 7. OK, so context 7 are two tools. One is resolved library ID and one more is get library docs. That's it for latest library docs. OK, so here I have.
And next JS latest docs. OK, so if you see it, it's using get library docs. So based on that docs it will be able to the latest latest, not the updated one. Huh. Correct. So that's important. OK, yes, so context 7 and file system.
OK so file system. If you see here it first uses in file system no no no. You have to expose like which directory you want to give access. So if you want to delete anything from desktop you can tell hey delete and this will last two times.
So as you see here read file is there, read text file is there. OK they have removed delete before it was there. So you have read file, read text file, read.

This is also read, write, edit, create, list, list. OK, that before it had. Do you think it's powerful if you give and then you the latest project with the this, this, this, this. Look at him.
And the way he has used MCVA, he has it. OK, so.
But tell the flow tell the flow OK create title and description.
Uh, it will fetch from Zira. Uh, first to back end, then front end. Hmm. The back end to the API is under key. Call postman and CP server. Hmm. Test value response. Yeah, so we'll back in complete and uh.
Designment. So data is then we'll complete, then Figma, Figma design copy, just copy the design. Figma frame ID, but you have to give that frame ID, yes, or would that be from the data?
The thing with from that we will not have. Yeah we will take the testing design by ID then front end ready and prompt me get back and say go ahead then play write and it will go into a testing. No it will open browser and you can see all these things.
So browser open some blogging blogging functionalities of test reports under your file system. We already had six.
I'm still is calmly listening. So generate. I mean sorry face generate. No we have code here.
Your code just test if you have code either test because everyone is using so we also wrote the code so the code is same. The only thing is when you use MCP as a tool. If you're using STDIO, you have to define it asynchronously. If you're doing SSE, you don't have to do it directly. It's like you define MCP tool equals to MCP tools.
And define that endpoint if it is what you call streamable. But here if you see it is not streamable and it is STD IO so you have to define a sync def run agent. Then this is your tool now so command is UVX.
Argument is MCP server quadrant ENV dictionary then timeout by default it is 30 but we have to keep it 50 as MCP tool. Now this MCP with tool will go here. So name is same model same tools is MCP tool instruction markdown.
And that's it. So here you have to use await and this is error. I didn't run it response dot content because the major code is inside tool calling because tool calling is getting index not the answer. So I didn't do response dot content. I want to see entire response.
And then query and run. That's it. Uh, do I need to run QVX? Because it's for me and regarding the QA. Oh, you have to download UVX.
UVX something code you will see this new OK so was our hybrid search then web

pages for initial now we have code.
No, my code will have all the chunks based on the output. This whatever you have right this entire code will be there and next time if you want to extract this code you can extract the code. So this code can be anything this code with Python rules.
Huh. You can copy paste, but the thing what we're doing is most of the time common templates. So then why for it? For example hybrid search binary quantization. Next time I tell hey give me the code for binary quantization, it should fetch it.
So I should not go to the documentation and search. So now what we do is how many ever cookbooks you have which is internal team push it entirely vector DB and then ask the MCP. Now what we have is.
MCV server. The more imaginative things you can do, the more output you can get on MCV. So you have cloud, right? That's our cloud settings, then developer edit config.
Here we have MCP server. So here we are using quadrant. So basically we have internal vector databases where you have URL and endpoint save. So now what do we do? One person will save all the cookbooks in vector DB. Next time that person, even if you don't check the cookbooks you can retrieve from.
What do you call quadrant database? I don't know how much people use this within our internal team, but it's very useful anywhere, anywhere, anywhere. But this is retrieving exact code rather than you going searching. Amarapas 20 cookbooks.
So in the 20 cookbooks, where will you go and search that code snippet? So if you could need a small snippet, it is there in the vector you you will get that back only that part. No, the entire code with metadata.
So there are three MCP server. I usually suggest context 7 context 7 one second.
No, both perfectly fine. I mean single page application. Multipage is not fine, but what you can do is you can use Jira. Jira will definitely be helpful because you are using it internally. But apart from that three common which will help in live coding is context 7.
Sequential thinking and any memory or retriever based MCP which is open memory MCP or quadrant. So this is retriever based, this is memory based. Now we I did the quadrant on my local docker. It won't have the MCP on local.
No, it's there. So if you check the documentation here, you could check OK, where is the documentation quadrant URL is HTTP local 633.
Then you have collection name embedding model. Then you can use this.

So you don't even need cloud if you need that, right?
We just need the one problem. So according to our end they can you experiment and explore so much and the people are just creating so creative. That is what he told us. There are more MCP servers. Usage is less but you have more MCP server. Now you see they're very developer. Developers will use it.
MCP is a developer tool. It's never a consumer tool. It's so I I kept memes on this. This is King.
Yeah, when either I auto complete my.
King the name. So now one king. Uh, this one. King of Baldwin.
So if you just start King Aldwin and then Ora, you'll get that image. You you have this King Balwin and you have explained my meme.com. It will explain what the meme is for. OK, I just type King Balwin and then Ora. There is a there is a you will explain my meme.
Explain uh my meme. Yeah. And then you will have know your meme and you'll just add a you know, uh, whatever you like in for.
Single link.
You will have all the meetings that goes and that is this is the one silence. Yeah, right on second one. Yeah, this is one of the same.
So it's so no, no. But this is for aura. It's like if anyone need any aura for me, it'll be like, it's a good person. I add this 4040 there like the little box.
That's like, yeah, I'll just copy this.
So if you take create your own NCP server, your cloud web, it has hmm an SDK ready for NCP server only do three lines of code and you can connect with anything you want. You can read our own tools.
But it only said I created when it's pretty easy to do very easy. So if you see now this is MCP client, so is a client, Claude is a client. Whatever we did now is a client, so we wrote our own code. But what did we do here?
What is this MCP server now? This one? This MCP server is something that you have written by what do you call?
Existing ones. But if you want to create your own, it's very simple. Just define your own tool. Google search. Google search is a tool and wrap that inside our MCP server. That's it. That's it. So you can create your own MCP server. The only thing is you'll build multiple tools. Every tool you know is part of MCP server. So now let's suppose I want to create.
And a set of tools that everyone that you and only authenticate people can do. I can

create my own cloudflare and could be server. I'll just create a list all that tools over there and I'll ask everyone to use that NCP server instead because they all are connected with their own identities.
So I was waiting for this all to complete before we can give that set of tools for everyone.
It is that the other thing causes too bad. So this is the one. So if you see here you define your own function which has get alerts above that you're doing MCP tool. Now this is part of your server dot PY. That's it.
But this is mainly for developers, users very less you'll find, but if you can get users you can build many interesting things.
Let him do one workshop. No, you should. You should show the entire journey because Jira MCP server is very useful. That is then and then even moving on from jira.
To get projects GitHub also has this tools. GitHub also has MCP server so you can work with the issues and the project and then you can move it like once this is done checked you move the card automatically. But the data project has many problems with your team but it is solved.
And for front end you can have Figma.
Even we try to use it for this. So here what you can do is copy your figma design, go to your cursor, just paste that link and tell write the code for it. It will write the code. But for that it needs high focus on the and the coordinator will write properly. Yes, everything is going to be replaced now, but it is.
The latest video that is now and again, don't. Who can manage both humans and agents, huh?
That means we can also manage the customers. Yeah, what you want to main Karan, but that's the agent and everything. I was, I was having dreams about it. Actually, it was so much in my mind. I was having dreams about all the this because after going home I tried multiple things.
Before coming in the morning also I was trying multiple things and I went to and everything for the car one and but it's so exciting. I was having lightwares and dreams and this is one example we missed out yesterday. So you're defining line item, you're defining voice data, so line item.
Description, quantity, total price and again unit price and total price. Then you have invoice data. Now this is your main schema right? In this main schema you have vendor name, vendor address, vendor e-mail, client name, client address and line

items. So this line item is coming from this class.
So now you have the same code you have from agno dot media import agent, sorry image image. So this is pydantic for schema. Then you have agent, then you have Gemini. So this code is same. I'm loading my Gemini API key.
I'm defining what I need to extract from invoice. So this is for multimodal. So when I had multimodal in mind, this was the example I wanted to show. So you had class methods and then you have system prompt like what you need to extract.
And then you have instruction, then you define agent. Mm-hmm. So agent model equals to Gemini, then roll, then instruction system message. We did this for. Routing and also for team. So the syntax here is same. Hmm.
The entire syntax to layer is same and obviously when we have to use output parser by default it is right. Now here what you need to do is just define output schema equals to invoice data. Sure use SN mode to be true. OK so instead of markdown now you have OK use SN so far. This is how you for this.
So for Gemini I tried to do this, but it shows that Gemini is used for open AI for. Yeah, because it does not support Jason. No, it supports no. I got the output from. I got the error that Gemini does not support.
No, this will support. This is not Jason. This is very similar. No, I mean I'm not using this. No, you're not using this, but concept is very similar, not similar. You're pretty much same.
So you have this. We discuss what we have. No, that is not possible. OK.
Chunking and all you have to do for that because not you'll get same image every time you'll have different images. You have to remove the unwanted one, but this one is there and then you just have to extract the content. That's it.
Do you have this like in working? Yeah, yeah.
So if I have terminal Python this multi modality to be concluded under 15.
This uh the complete things should run under 15 minutes. It could it could not for 10,000 pages, but then you have. So even if you're doing agents and if they can you have Jason.
But like uh the we cannot be able to like know it you know can we split the one but see we have to write a some funny name address.
Parallel even get land graph. I want to I want to use the land graph and land. I land graph. I cannot use like the flow chart. Yeah because we are using.
OK, but yeah, vendor name, name, address, contact information, invoice details and line items is list and all the details. Line items, description, quantity, unit price, total

price.
Description quantity unit price. This is for multimodal that are defined. Yeah that I need score for one. This is like yeah yeah it will be less also, but Gemini or any LLM won't comment that we are less.
So far everything it has taken 0.99100% confidence. Now what this is good. This is on multimodal. We send an invoice you can set with agent set a reminder for the account team to be paid by this date. This is the amount and stuff. You can do everything now.
Agents like if you know how to use your tools, agent is very easy. The only thing is tools like tools. It is both what context you're getting, what action you're taking. Action is a tool, then context is a tool.
We are getting discovered.
Yeah, mostly that's it that that has to take care of.
Oh, you apply for OK, but it's not free. I think it's not free. OK, so this is instead of continue, we will start.
So this will be on. No, no. Oh, OK. So here I added continue. OK, so instead of continue, we'll start on 9th, which is tomorrow. OK.
OK.
But MCP is done.
Only when you are in demo, MCP is very good. In demos MCP is very good, but when we deploy one of my.
One of my I'm seeing a lot of comes to report creation. Report generation is easy with the like they're using the but.
Now we're using this, I guess if it is, if it is in MCP, it could be way more easier now. Yeah, but it needs to have. But if it has the APIs, if it has APIs, just copy the entire API schema, give it to Claude. It will write your entire MCP from scratch. Yeah, yeah, Claude can write MCP from scratch and very simple logic.
No, Claude is good with free CAD, free CAD. Who? Who do you think wrote the free CAD MCP? Claude, a person plus Claude. And you have one more software which is Solidworks. Solidworks. Yeah, Solidworks. What do you call? They don't have MCP server. We built it from scratch.
Zero code offline. We wrote by ourselves. If you open Solidworks, what do you call them? No Solidworks documentation. Solidworks. Solidworks is a free CAD is free. Solidworks is paid. No, no. So Solidworks doesn't.
They're MCP. We already have. So tell me. We have MCP servers for field month, so

for wire framing purpose if we have the right BRD in place, it will create.
No, it cannot create a Sigma. It can only face from face from. It's not clear. No, no, right. So now if you see this is the SolidWorks documentation line by line we started giving a hey, see you will. I'll give you all the documentation.
This is my requirement. If you think this is useful, tell us or no. So for whichever documentation it told us, yes, we saved that URL and then we started writing code. So but from with cloud, like how would you give away this much context? And so only I'm starting with one documentation page, right?
I give requirement. I give the documentation this day and I'll tell hey tell yes or no if this makes sense with the requirement. OK then for which one it's a yes. Then we give the documentation and tell it to write code. So it will write NCP tools just for one and you just integrate it. No, just yes or no. Starting is just yes or no. Yes or no. Then you figure it out. OK these are the these are.
Documentation. Then I start giving documentation and tell to build that tool.
But it is a low contextion cloud has low context. But I'm creating different chats and then building it. OK, so different chats, different chats. So you will then, you know, merge it hard together. OK, merging of tools is my part. OK, but writing the tools logic, I don't know what this is not Python.
This is C# something so I told this is the documentation. Give me Python and it gave me Python code.
OK.
The rules, the responsibilities. So was this recorded? It's saying waiting to other switch on. Yeah, yeah, it was recorded. So if you see on more, click on more and then you will see uh.