Transcript of Video

Introduction

Eager to be part of the

rapidly growing AI field. Then you are in the right place. Today we are exploring the forefront of

technology with generative AI, a field that transform how we create and interact with digital content.

We'll start by breaking down the basics including what generative AI is, how generative AI model function,

and what makes them so powerful. From there we''ll look at the advantages

of generative AI and explore what the future holds for this transformative technology.

But it's not just about the tech. We'll also examine the ethical

consideration to ensure

you understand the broader implications. But wait, there is more. Next, we'll introduce you to

large language models.

That is LLMs, featuring

popular ones like GPTs Claude, and Gemini.

And that's not all you will get hands-on with prompting techniques and learn how to create your

very own LLM app for Android. It's easier than it sounds. We'll also explore top AI tools

like ChatGPT, and GitHub Copilot and show you various demos on how

to use these tools to cap it all off we'll explore advanced concepts like Langchain and RAG

with compelling case studies to wrap things up. By the end of the video,

you will have the skills needed to build advanced AI applications. Continue watching to excel in generative AI

and check out the time stamp in the description to find your favorite parts.

Introduction to Generative AI

Let me make you understand with a very simple example what is generative AI.

Transport yourself back to your childhood, you had a lot and a lot of

toys to play with. You would keep the toys in one box. Now, also imagine that

if you wanted some toy which is different you would not get in the market.

But what if I tell you that this box is a magical box.

And if you input your understanding of what you want

in your new toy with instructions, it can create a new toy for you,
which is not available in the market. Now, this toy can be a bear with
unicorn features and wings. What if it generates for you? This magical box generates a toy
which is very unique for you. This magical box is nothing but generative AI.
Generative AI actually is not a magic it's a fast and rapidly evolving
artificial intelligence system
which creates, generates, transforms, content that can be
text, video, audio, image, etc. based on your input.
So if you want to understand it technically generative AI or Gen AI
functions by employing a neural network to analyze data patterns
and generates new content based on those patterns. Neural networks are nothing but
a mimicry or a replication of your biological neuron based on how it gets from brain
the activity from brain,
and you do your work. It's nothing but a mimicry of that.
Based on that mimicry,
it analyzes data patterns and generates new content for you.
Let's now quickly see what
is the difference between discriminative and generative AI?
Suppose you have a dataset of different images of dogs, cats,
you provide this as an input to your discriminative AI, which acts like a judge,
and it classifies all this into a set of images
between cats and dogs.
This is discriminative AI. It classifies. Now let's understand what is generative AI.
You have the similar set of cats and dogs, but now your generative AI
is acting like an artist.
It creates a new species of dogs for you. That's why generative AI is nothing but
AI system that transform,
creates, generates your own content based on your
instructions, like an artist. Now that you have understood
what is discriminative AI

and what is generative AI and what is the difference between the two. Let's understand why is generative AI

or Gen AI trending. Gen AI has impacted various fields,

be it text, audio, video, any input and those inputs in various domains

like data, management,

tech, health care, and entertainment. It has creative applications

such as Dall-E, ChatGPT,

where you can input what you want

and get output from it.

For example, if you want to create

an image what you think or perceive

as a concept and you want it,

you give a prompt for your generative A.I.

model. And if you create that image for you, so your input is a text,

but your output is an image.

That's why it's trending. It does not depend how traditional A.I. is dependent

on what form of input you give.

The same form would be output. However, Gen A.I. works on your inputs,

on your instructions.

That's why it's trending. It is impacting a lot of things. Be it create a field, be

it research field, be it business.

Professionals are using tools

like Chad Djibouti to create or generate code

so that they can create something new.

The researchers are actually developing

new and new large language models

based on which we can create

new generative models and can do new a new task

each and every day.

That's why generative A.I. is evolving rapidly and that's

why is close to magic for everyone. Now that you have understood

why it is trending now, let's understand
how it works. We give an input to generative models.
Gen AI works on generative models. We give an input. It can be text, audio, video, any format.
Those generative models
are then retained on the data
and they are fine
tuned to do the task that you want. It can be fake summarization,
it can be sentiment analysis,
it can be made generation,
it can be audio generation for your YouTube channel
or analyzing your customer feedback.
If you are a brand or marketing firm,
it can create goods, whatever you want.
You give a prompt
what you want, explaining it that what you want, and it's fine
tuned and gives you that task.
So this is how, in a nutshell,
generative A.I. model works. So now let's see,
what are the different types of generative
first one is generative
adversarial Network Gans. It's a type of AI.
Where do models one generating the content
and one judging it work together
to produce realistic new data. Second is variation, auto and goodness.
This AI learns to recreate and generate new similar data to what is Transformers.
Transformers is an AI which learns to produce sequences using context.
Fourth is diffusion model,
which generates data by refining noisy starting
until it looks realistic.
Now that you have understood
what are the different types of generative AI, let's quickly look through different

applications of generative Air Force.
One is content generation. It creates and generates
whatever textual or
any code that you want customer
support and engagement. If you are brand form, it helps you
with that data analysis and data science.
It helps with visualization,
it helps with analyzing any data, it beat it any data you want.
You are a brand form
or you are a technology firm. It will help you analyze your data
and create
new automated tasks for you. Or it would create new perception
for you to take over.
Then it is go
generation and software development. We have research and information
retrieval as well.
There. It helps different researchers. It helps different professionals
to grow and retrieve,
extract information required
from different or reduce data sources. Then we have machine translation.
If you are a person who do not understand
a language and you are watching something or reading something which is in different
language, you can use generative models
to translate text or audio or anything into the language
that you require.
Then we have sentiment analysis
which actually takes feedbacks or any text that you have to give you.
Is it a positive,
negative or neutral sentiment and so that you can analyze
and take decisive decisions?
Other domains here include healthcare,

transport, everywhere. It helps generative models.
Generative A.I.
is helping each
and every domain in their perspective, how they are applying this
technology change in their domain. Step into the fascinating world
of generating a technology
that is transforming industries
and creativity alike will break down. What they do is explore the advantage
it brings and discuss
its potential for the future. Additionally,
we will explain how generative models work
and address the key
ethical considerations that come with her. Amazing, this powerful technology.
So let's get it started.
Advantages of Generative AI
Let's now understand the advantages
of generating air tended to offer numerous benefits
across various
industries and enhances
creativity by allowing tools like Deli to create unique artwork
and use that to compose original music.
Does expanding the creative possibilities
for artists in fashion and design tools like Runway
significantly increase efficient C
by speeding up the creation process?
Saving both time and cost personalization is another key advantage,
as it can generate tailored
content for marketing. Improving user engagement and conversion
rates in the field of drug discovery helps
accelerate the development of new treatments
by generating novel molecule structures.

Additionally, AI's ability to hand in large scale content generation,
such as product design
or media, allow businesses
to adapt quickly to market demands
overall and reduce is driving innovation, creativity
and efficiency across multiple sectors.
Now let's
talk about the future of generative, which holds exciting possibilities.
The Future of Generative AI
We can expect A.I. to increasingly collaborate
with artists and creators,
pushing the boundaries of creativity
by coauthoring works and composing music.
As a technology advances, it will offer even more personalized experience
settings tailored educational materials
that adapt to individ three
learning styles integration with augmented reality
that is, air
and virtual reality
that is, VR will enhances this environment by creating more immersive
and interactive experiences.
Ethical consideration
will be a major focus with a force dedicated to what
and sharing responsibilities,
use and addressing issues
like data, privacy and authenticity. Furthermore, dendritic will drive
innovation across various industries,
including health care and finance,
reshaping how we approach and so on.
Complex problems. Let us understand about
the ethical consideration engendered.

Ethical Considerations in Generative AI

Do it because as generative technology evolves, several ethical consideration

must be addressed.

One major concern

in the potential for A.I. to create real realistic,

but misleading content

such as deepfakes

which can contribute to misinformation. There is also the risk of

perpetrating biases present

in the training data,

which can impact fairness and accuracy. Ownership of A.I.

generated content poses

another challenge as it is important to establish clear guidelines

for intellectual property rights.

Additionally, a AI system

might in Edmonton lead reveal sensitive information from the training data,

necessitating

strong privacy measures. Finally, it is crucial to establish

standards to ensure that gender

is used ethically and responsibly, avoiding misuse and harmful applications.

Now that we have covered the issues of

gender, let me ask you a brief question.

Which aspect of identity

are you most curious about and why? Let me know your responses in the comments

section.

Next, we'll explore the exciting

realm of large language models. That is excellent.

We'll begin with an introduction

to analytics and take a closer look at some of the most popular ones,

including Deep

Cloud, 3.5, Sony and Gemini. We'll also delve into opening a zippy DBA

and guide you through the process
of creating an excellent app for Android. Ready to Harness the Power of HoloLens.
Let's begin. What exactly is an app? Imagine you have a super smart friend
Introduction and Phases to LLMs
who has read every book,
article and blog post in the world.
This friend can chat with you
about almost anything, help you with your homework,
write stories, and even tell you jokes.
That's pretty much what an L.M. That's in technical terms, and L.M.
is a type of artificial intelligence
that has been trained on vast amounts of data
It can understand and generate
humanlike text based on the pattern
it has learned from all that reading.
This is part of a broader
field by generative A.I., which focuses on creating content
that resembles human output. Let's now break down the basics
logic, language and model. Logic just means that the model
has been trained
with a massive amount of data. We are talking about billions of words
from books,
websites, articles and more. Model refers to a set of instructions
or algorithms that help the air understand and generate language.
Think of it like a recipe that guides how to make a cake.
In this case, it is guiding the A.I. on how to create meaningful text.
Let's talk about language. This is all about words, sentences,
and how we communicate. When you text your friends, write an email
or post on social media,
you are using language. So now that you understand
what each word in LLN stands for, let's quickly

see what is the technical definition?
Elements are language models made up of neural network
with billions of parameters that
are trained by Self-Supervised
Learning on was the amount of unlabeled text.
Now let's dive into
how do L'alliance work? Let's simplify this with an example.
Imagine you type the sky is dash
dash Dutch. The NLM might predict
the next word to be blue
because it has seen that phrase so many times during its training.
Large language models like DPD
Ford operate based on complex neural networks,
trained on vast amounts of text data. Here is a simplified explanation
of how they work in two phases.
The first phase is the training phase
followed by the inference phase. The training phase has four steps.
Step one data collection that is gathering diverse
text data from various sources.
Step two pre-processing, which means cleaning and tokenizing
text data into numerical representations. Step three Model architecture,
which is nothing but designing
the neural networks structure. Typically, a transformer step
for training is adjusting model parameters by predicting the next word
infant density.
By this training phase,
the elements are being trained on what amount of data to do.
Further predictions. Now coming to the next phase,
which is inference phase.
It also has four steps. Step one input processing.
That means tokenizing and converting

the input text into embeddings.
Step two Generating output. Now using the model to predict
and generate the next what?
Step three Sampling which is selecting
words from predicted probability.
Distribute. Step four Last
but not least, post processing,
which means converting
generated tokens back to readable text. After the training phase
in the inference phase,
the elements will recognize the pattern sample content and give predictions.
Apart from these steps,
there are three key concepts
which we should know
to demystify elements. First, attention mechanism.
This allows the model to focus on relevant
part of the input text.
Improving its understanding
of context and meaning.
Second embeddings which
are numerical representation of words
or tokens that capture their meaning
and relationship.
Third Transformers. Transformers are the architecture
that uses self attention to process
input data in parallel,
making it efficient and more powerful. Let's see some examples of elements
in action.
Example one Charting your plant can be
what is the capital of Japan?
llN replies The capital of Japan is Tokyo. LLN knows this because it has read

a lot of information
about countries and capitals. Let's see Example two About writing
your project can be
Can you help me write a story? llN says, Sure.
Once upon a time in a land. Finally there was a brave knight
who marched.
The NLM can generate creative text based on what it has
learned from reading stories.
Example three Lets see answering questions Your prompt can be explained
photosynthesis in simple terms.
Replies Photosynthesis is a process
by which plants use sunlight to make food from carbon
dioxide and water.
It simplifies complex topics
by using the patterns it has learned from education
news content.
There are different types
of elements, each designed for specific doubt and application.
Now let's explore these types. First are base models.
They are trained on a wide range of data
and can perform general language
understanding and generation dots.
The audience specialized
for any particular type of instructions,
but are versatile. For example, Deep three,
which is known for its versatility
and ability to generate coherent
and contextually relevant text.
Another type is instruction based models. They are fine tuned to follow
specific instructions
better than base models. They can understand and execute tasks

based on detailed prompts
given by the user. For example D five
that is text to text transfer transformer, it converts all an LP problems
into a text or text format,
making it highly versatile for tasks like translation,
summarization and question answering. Another example is instruct DPD,
which is tailored to follow specific user
instructions better and provide
more useful and safe responses.
So these were the types of elements. However, there are paid and open
source elements as well in the market. Based on your usage and budget,
you can explore them.
Some paid elements are good for for all Microsoft
Azure, Openai services, etc.
and open source are Jeopardy, Neo,
but etc.. Let's see.
How are excellent reviews
ionizing the world of air in education?
And they can help explain
difficult concepts, answer questions, and even do to students in content
creation.
The assist writers, marketers, creators
by creating generating ideas,
drafting content and editing,
encoding the assist professionals
with generating code to their specific problems
in customer support.
Many companies
use elements to bother their chat bot providing quick and accurate responses
to their customer inquiries. Apart from all these fields,
elements are being utilized

in every field and domain, but it's not all smooth sailing.
There are challenges too. Let's see one such challenge,
which is bias.
Since Elam's learned from human
written text, they can pick up and reproduce
biases present in the data. Another one is misinformation.
They can sometimes provide incorrect
or misleading information, especially if the data they were trained
on was flawed.
However, keeping challenges aside, the future of excellence is exciting
as they continue to improve.
We can expect them to become
even more accurate, reliable and versatile.
They will be better at understanding
context, handling complex tasks, and even learning
from smaller amounts of data.
So there you have it. Large language
models and ends are like super
smart friends, followed by tons of data
and advanced algorithms.
They can chat right? And help us in countless ways
as they continue to evolve.
They are set to become even more
integral in our daily lives.
As you know, the large language
models are advanced a system designed to understand and generate human
like best.
Let's talk about
some of the most popular elements. First is deep differ from open
Xs in generating coherent
and relevant text used in chat bot,

virtual assistants and content creation. Then bot created by Google
understand word context
by analyzing surrounding words, making it effective for question,
answering and sentiment analysis.
And also P5 also from Google handles
diverse tasks
such as translation and summarization
by treating them as text to text problems.
After that blog, developed by Anthropic, focuses on providing ethical
and safe response is making it suitable
for sensitive applications.
Next is Llama, which is from ETA, offers high performance in various language tasks
providing useful
for research and Next-Generation
Lot 3.5 stone. It is an advanced
version of plot emphasizing
improvements in safety and effectiveness
for Omni is a smaller variant
of CBT four designed for efficiency
in various application. Gemini from Google is known
for its advanced
capabilities in handling complex language
task. These elements are central
to advancing natural language processing
and driving innovation across many fields
as these models continue to evolve.
They will play
an increasingly crucial role in shaping the future
of AI and its applications.
Introduction to OpenAPI GPT API
Let's understand Introduction to Open Air API how this particular thing works

and what does open air.

What does open air? It's a company that it will cater

in order to work with chat bots

generated via applications, different kinds of models, elements, etc..

Basically it is dealing complete artificial intelligence domain,

which is booming nowadays.

Open Air has a platform

where you can generate the API keys

and you can integrate those

into your applications.

API features. What are the features? It will cater for next generation

completion and conversation capabilities.

So talking about the next generation, it is always dealing with

providing a new text which is not in your imagination.

But one small question. See, I want to point three on so and so.

It will give you a complete poetry that it is not plagiarized.

It has been trained in that level. It can think about writing poetry.

It has lots and lots of data behind how it is dealing with that.

What is to be categorized

There comes classification, summarization and many other machine

learning and data science models,

artificial intelligence models. That is giving you the answer

for your next completion.

If you give a prompt in that incomplete way,

it will try to complete that.

If you give it a spelling mistake, it will

collect your spelling and ask you back.

Was this your idea? Do you want to such sequential thing?

It will question you back

in our interactive conversation of

what we do with you. It's the conversation

how it will answer us with the help of already available data.

It has been trained on and it is updated

every time you talk to it. That's how the model works.

So these are the new features we have. Next comes to fine tuning

and customization for a specific task.

Say you are building certain model which

has been integrated to your application.

You are using Openai platform. You can generate your own model. It can cater to your own set of questions.

Say, for example, chatbots

and now some other other shopping websites or jewelry shop

websites.

It will try to ask you what you want the launch of chatbots,

which will address and also will help

certain percentage of customer

care services without human intervention,

which can be done with the help of machine

100%, it will be solved. Other aspects it cannot.

A very good example for this is Swiggy. You can do a set of questions

which is already present.

Where is my order?

Delivery guy is not moving. So when you put this,

my order is getting delayed.

It will give us a set of answer. It is already that. What is the current status? Still, if you are not convinced by the bot

answer, you can go for the agent.

You can talk to a human

being that they'll interact. They'll call the delivery guy and ask

what is the situation and update

or something like that might happen. Before

introducing your particular chat directly

to the agent, they'll try to solve

with the help of what that means. We are trying to reduce the work
which is put upon humans.
We are using the technology at the same. This is the best example for the feature
which is currently in use in the apps
which we use in our day to day life.
How do we get that with this API? We have to just log on to open a website,
create your account, sign up, or if you already have an account sign
in login,
generate an API key and keep it. Why you have to generate the API key.
I'll let you fill $5 of content. You have to be in order to improvise
the API key
in order to improvise your API key usage. You want to know
more about what is open AI?
How does it help for GP, the API? Everything. You can just go to the official
API documentation and understand
more about this. Now let's understand how do we generate
a open API key for that? You have to go to Google type
open EIA log in.
Once you click on logging, if you have already logged,
then you will get two options. One is to go to Chargeability.
Another one is to bar API. You click on API. Once you click on the API, this is how
your open API platform look like.
You could see a menu here to your left
stating API keys. If you click on that,
it will launch API keys.
Before that, I would like to tell you
I was talking about the GPT models.
Right? So these are the models available for now. GPT 3.5 00125w106 and 16.
These are the models. You can select the model
and you can work on it. Let's go back to API keys
and click on keys because this is how the API keys

generation look like.

And if you want to create a new API key, click on Create your secret key

and you can name the particular one. I'm naming it as demo.

You can also give the restrictions

if you have to control certain things. It can be of only the static

or all, just like the share option

you have in your Google drive

or your Google account. So create a secret key

and it will generate and display

the secret that you can copy and paste it

in one particular notepad

so that you can use it again and again.

It is taking certain

time to generate the key. Once it is done, it will display and you

will also have an option code copy for it.

Here we are. Each date app is generated

and you can copy the key.

You can press as then see, you have to save the secret key

somewhere

because it won't be viewed again due to security reasons.

That's right. You have to keep it discreet

and noted in that notepad separately. You can get it back again if you want to.

Again, you have to create that new API. You cannot copy this complete API key.

Again, the creative

API key is listed in the dragon. You have the options to edit the key.

You can just change the name

and permission. Nothing is you don't have access to. Again, copy the complete E

and you can also delete that existing key.

This is how the API key page

look like in opening a platform book.

You are clear how to generate this

and save it in a place and use it for your coding.

Today we are going to explore a truly transformative tool

for developers of all levels.

Claude Sonnet 3.5

Your new AI powered coding companion.

Imagine having a super smart friend who can handle all the tedious parts

of coding for you,

allowing you to focus on being creative

and solving interesting problems.

That's exactly what Claude Sonnet

3.5 does.

Claude 3.5 Sonnet

What exactly is Claude? Sonnet 3.5? Claude Sonnet 3.5,

which came out on June 21, 2024,

is the newest and most advanced

UI tool from a company called Entropic.

This assistant is made to help people

who write computer code called developers by making their work

easier and faster.

Claude Sonnet

It is like a super smart friend who can understand what you see

in everyday language.

Instead of typing complex codes,

you can just tell Claude what you need in plain English

and it will write the code for you.

For example, if you need a program

to add numbers, you can simply say, I need a program to add numbers

and Claude will create it for you.

This new tool is not just about

making things easier. It works

faster and has better safety features

to protect your work and data. Claude 3.5 Sonnet helps
everyone from beginners to experts
to write code in a way
that is more natural and less complicated.
Claude 3.5 Sonnet changes how we write code by making it simpler, safer,
and more accessible
for everyone,
no matter how experienced they are. Alright, let's talk about the different
versions of Claude Sonnet 3.5.
Imagine you are picking a tool
that fits your needs perfectly. Whether you are just starting to code
or your already approved
first step, we have a free version. This version is fantastic
for beginners and hobbyists.
It's like having a helpful friend
who can write basic code and fix simple box for you.
You can access this free version
on Claude Dati or even on Claude iOS app.
It's a great way
to see how we can make coding easier. Now, if you're looking for more power
and features, that is a paid version.
This version is for those
who need to tackle more complex projects
with the paid version.
You can generate more advance
code, get better debugging, help and integrate with more development tools.
If you are really into coding,
this might be what you need. There are also special
plans like Cloud Pro and the team plan.
This gives you a lot more power to work
with. You get higher rate limits,

which means you can do
a lot more coding
without hitting any limits. And that's not all.
Claude Sonnet 3.5 is also available
through some major platforms
like the Entropic API, Amazon Bedrock and Cloud Vertex.
So if you are using any of these services, you can easily integrate
Cloud Sonnet into your workflow.
We discussed about the cloud sonnet paid
and free version, right? So there are different versions
of your cloud sonnet, as you can
see on the graph here,
which is between your intelligence benchmark score
and your cost price per million tokens.
So these versions, we will see
how do they perform or what are the capabilities they have in
just a brief out of these versions.
The first one you can see is your Cloud
three haiku, which is on the lower end of your intelligence
as well as the cost.
This is actually the basic version
of your cloud where you can do some simple tasks as well as you can see
that it is for your beginners, right?
And then you can see here
that you have cloud three sonnet,
which is an intermediate tree
between your intelligence benchmarks
and your cost price per million tokens.
Right. So this Cloud three sonnet
was used for performing complex tasks.
And here you can see that

you have three opus, which is on the higher scale of your intelligence
as well as on your cost.
Right. So this was a version
which was being used for performing complex tasks as well as used
for your profession in works to be done.
Now you can see that we have Claude 3.5 Sonnet, which is on the higher
end of your intelligence benchmark
as well as on the cost price
per million tokens. So we can say that this class sonnet 3.5
is one of the advanced version of cloud. You will be able to perform
your complex dos
and as well as it will help you
in all your debugging. And you just have to write your queries
in your natural language
and it will give you the outcome of it. Right? So this is about your versions
of your plots on it.
Now, coming to the features of your Cloud
3.5 Sonnet, let us see what entropy
latest breakthrough is, right?
Which is actually making waves
in your artificial intelligence community. So we will see what are the new features
and the new enhancements
which cloud 3.5 Sonnet has. Right. So for the first one,
what it is discussing here is industry
leading performance, right? So you can see that 2:00, 3.5 sonnet
sets a new benchmark
for a performance, Right. You can also see that it is outperforming
its predecessors and the competitors.
Right. So what are the competitors? You have your open your eyes GP photo and
you have your Googles, Gemini, 1.5 Pro.
So you can say that your cloud Sonic 3.5

is actually giving you better performance
from the other open
eyes like your GP photo or your Google is Gemini
1.5 growth, right?
And then it says that these advancements
are significant, far exceeding the capabilities of Cloud
three Opus,
which has the higher intelligences
as well as the cost. We just discussed that here.
Right now
the next feature is enhanced speed, right?
So this enhanced speed is twice
that of your Cloud three opus,
which we just saw in the graph
that it is on the higher end of your intelligence
as well as on the cost, Right.
The increased processing speed facilitates
handling complex tasks.
As we said, that cloud 3.5 SONET
is capable of handling complex tasks and multi-step
workflows more effectively,
which will help in opening
new possibilities for real time applications
such as your finance and your health kit.
And indeed, a good feature
you can see that it is providing you increased and improved efficiency
with enhanced speed.
The next one is advanced
coding capabilities. Right. So Cloud 3.5 SONIC stands out
for its advanced
coding capabilities. Right. This is what we said

that you just give something
in your natural language,
like, say, for example, you want to write a program on printing
Hello World.
That's a common example which everyone
takes in any of the programing languages when you just want to start off with,
right?
You're just telling that
it will give you the complex tasks.
Maybe you're trying to build a website,
maybe you're trying to integrate your website
with the frontend and the backend,
or you're having a full fledged database. You're trying to have something related
to e-commerce or you're having some tasks
which is appropriate for those projects
which you're trying to work for. Right. So It will give you these advanced
coding capabilities.
So what does the internal evaluation tell? The internal evaluation tells
that it has solved
64% of coding problems. That is an improvement,
over 38% solved by your cloud free opus.
Right. This makes it a powerful tool
for software development. I'm code maintenance, so 64% of coding
problems is being solved by your code 3.5. So net, which makes it a remarkable tool
for your software development
and or maintenance. Right. It's ability to independently write, edit
and execute code
coupled with sophisticated reasoning. We will also see a simple demo on the
logical reasoning which your clock 3.5.
So on. It would provide for some of our prompts you can see okay and
allows it to handle complex coding tasks

and code base migrations efficiently. Now coming to the next feature,
that is your super visual reasoning.
Right. So what does your clock 3.5 solid does
with your visual reasoning?
It is trying to interpret your charts, graphs and complex backgrounds,
but it accurately transcribe
text from imperfect images
which are crucial for industries like retail, logistics
and financial services.
One of the cool features of your clock 3.5 sonnet is superior visual reasoning,
wherein it is
trying to extract
the information of the visual data, even when the image quality
is such a cool feature, right?
So let's move on to the next one. That is your innovative interaction
and artifacts.
Artifacts is one of the new features which was introduced in your Claude
3.5 sonnet.
So what exactly is this artifact? This artifact
transforms
your cloud from an conversational AI into a collaborative work environment.
When users generate content like code
snippets, fakes documents, or web designs, these artifacts
appear in a dedicated window,
allowing real time
editing and integration into the projects. So we will have a slight demo on
what exactly
are these artifact new features
which cloud 3.5. So on It does in a little way. Right.
So let's move on to the next feature. That is your cost effective accessibility
as we saw that glow 3.5.

So it is higher in your intelligence
as well as it's in the intermediate
three for your cost.
Right. So your cloud 3.5 sonnet is accessible for free on Claudia and Claude iOS app.
We have already discussed that. Right. And you have this limits for pro and team
planner subscribers, right?
So what's the next one? That is commitment
to security and privacy.
This is actually the important feature
with your cloud 3.5. So it provides
because security and privacy
is much crucial
because you are relying on your A.I.
platforms and then you are trying to
or have your task being accomplished and completed by it.
Obviously, you would want to have
your security and privacy being maintained and not being exposed.
So what exactly does these security and privacy features
have?
It's like it says that Anthropic
has prioritized security and privacy with Claude 3.5 sonnets.
Right. Okay. Has maintained an ACE L2 rating
and even external experts,
including your UK's Air Safety Institute, have evaluated the security mechanisms.
Right. So we can trust blindly on your Claude 3.5
sonnet. But. But watch out.
What are you trying to have your task
being completed? See to it that still
the personal informations are not
being given that right. So that's about your security and privacy. Now next, moving to your part
of growing a family, right?

Your 3.5 sonnet is a part of a broader AI model lineup
as we saw the different versions.
Like you have your cloud hycu, you have your cloud opus
as well as you have
your cloud trees on it, right? I know 3.5 on
it is the higher end version of your plot.
Right. And the next feature it has is enterprise
focused design.
Right. So it will help you in handling complex workflows and integrates
with existing business applications. Its contextual understanding
and interpretation makes it ideal for
tasks like customer support
market analysis and data interpretation.
Cool. Right. We are not cloud. Sonic not only helps the big nurse
for the professional use,
but it also has a focus on
your enterprise tasks as well.
Like you have your customer support,
market analysis and your data interpretation, and the next feature
is user driven development.
Right. Your anthropic values user feedback as a crucial component of your cloud 3.5
Sonnet development.
Right. It is great right that they are valuing
the user's feedback across the globe.
Right. And 3.5 Sonnet redefines the capabilities with enhanced intelligence
speed and advanced features.
We discussed what that one features
like we have your Artifacts, which is one of the most recent feature
which was added in your Cloud 3.5 sonnet
and we will try to see a smaller demo on how these artifacts work in a while now.
Now next, moving to the advantages

of Cloud 3.5 Sonnet. Why a game changer?
Superior performance and cost efficiency. We already saw the different features
it has the cost effectiveness,
the visual reasoning capability tests. Right. And the capability of trying to handle
complex does enterprise driven methods.
All these are some of the features
which your cloud has. So based on these features,
what are the advantages which we have
as we know that it is your MLP
capabilities, your Claude AI is
and advanced has an advanced
and LP capabilities as well as it is combining with your cost
effective pricing intelligence
with less cost is what we clocked 3.5. So notice providing yields right
its ability to grasp nuance.
I'm human and generate high quality
natural content makes it versatile
tool across various applications.
Right. Advanced coding proficiency. This is one of the significant advantages
wherein
it is trying to solve
your coding problems. It is also fixing bugs
so nice and also add functionalities
to your open source code base with ease. Right. Okay. This makes it particularly effective
for updating legacy applications
and migrating code bases, providing
a robust solution for developers. So basically your clock 3.5 sonnet
is having that advanced coding capability,
which will not only help the beginners,
but also the people who are trying to work
on different softwares and different programing languages

on some complex tasks as well.
Right? So the next ADD one page is enhanced
revision capabilities. We saw that
one of the features of your cloud sonnet
is the reasoning, the visual reasoning,
which it interprets, right? Even when the image quality is less
if is trying to interpret
the exact information of the image. Right. This is one on that one page
that is enhanced vision capability.
Read in the improved version of your 3.5
plot. Sonnet has the ability to interpret
and analyze your visual data, right?
We also saw the innovative features
of your artifact, and I said that we will try to see the demo in a white light.
We did see the feature of your artifacts,
right? And I even told you that we will quickly
see a demo in a white cloud.
Claude Artifacts
Artifacts are designed
to make your interactions with outputs more dynamic and versatile.
Whether you're working on with your code,
data visualization or text artifacts
allows you to directly manipulate
and enhance your outputs
within the cloud interface. Cloud 3.5.
So it offers limited and free access. But for regular use,
a pro subscription is recommended.
The Pro subscription provides unlimited
access to features like artifacts. You can access Cloud 3.5
Sonet through Google Cloud, Vertex
II and Amazon Bedrock with usage

based pricing.
We have been listening
from a way through that artifact. This one of the new feature which was
provided by your plot 2.5 So in it.
So as you see that I locked in myself to my cloud dot
I right, you can give your email ID
Demo on Claude Artifacts
and create a password and try to log in
even with your Google account as well. So we will try to explore
the artifact features here
so you can see that
I'm going to my profile here. Right. And then feature preview
when I click on Feature Preview,
it is giving me Artifacts preview and provides feedback
on upcoming enhancements of your platform.
So that's what the feature preview is. And here is what artifact does that does
ask cloud to generate content, Little code
snippets, text documents or website
designs, and cloud will create an artifact that appears in a dedicated window
alongside your conversation.
Initially it is also
let me just any bullet, right? So after I have enabled,
I'm closing this window.
So let me see that. I'm just going to try out a simple tic tac toe game.
You can see. You can see. Right. Okay. So I will just tell that create a
tic tac toe game for two players and
provide me with the sample
output and see Python language.
Okay, So now let me just click on go here.
Wow. Such a cool explanation on the working of your tic tac
toe game as well.

Right.

This is what is your artifact. Right? But and it is trying to give you the code

as well as trying

to give you the air capability t

with the help of your artifacts. Right. So let us see

what is the output that is giving you?

It says that I would create a simple

rectangle game for two players in Python. So that's what the requirement

that I had given for.

Okay, so click to open the code. You can as well download your code here. The implementation is as follows.

It says Springboard displays

the current state of board check when it has some functions there.

And then it also says

how to run the program, right? How to run the game. Run the game. You can save this

code and in Python, fine.

I'll just give you a quick tip. You can go to your online editor replayed and try to open a Python project

and you can just copy this code

and try to check

whether the game is working or not.

Right. As in the prompt. I also told

that I need to have a simple output of how the game might look,

so you can see that it is providing me

on player X, player

zero, Player one and so and so forth. Right. Okay.

So is it all that your cloud artifact does

or does it have another capability or more capabilities

as a test promised?

What? So what will we do is we will just try to

write another prompt saying that we see

I want all presentation

for a class wherein I am
teaching this game. So provide me with
all presentation
starting from agenda Introduction
followed by the summary. So let's see what it gives.
Wow. It's creating a presentation for me
and you can see that the presentation is on the artifact
window here and it is giving me
exactly what has to be done
and what has to be conveyed when I'm trying to teach about this
detector game in a classroom.
All right. So let's see, what does giving.
Hello. So what did you give me this
presentation outline provide
structured approach to teaching the tic
tac toe game implementation in Python. It covers game basics, great Python
implementation, detailed code
walkthrough and suggestions
for further enhancements. Great and key points of presentation
that starting with Introduction Basic
Strategies Implementation Code,
Walkthrough Demo Possible enhancements also it does giving.
I never asked any possible enhancements
for the learning and creativity
well in question and answer session.
That includes to address your question
and answers and encourage discussions. Right. Such a great thing
it is trying to give me.
So let's
see what it has added in my presentation. Right.
So as I said, I wanted an agenda grade.

It is giving me an agenda and introduction
because asking me to give a brief history of the game,
popularity and educational value.
Why are we implementing Python? So I and so forth. Right. So a perfect piece of information
which I can just take this
and go to my classroom and have a discussion on my tic
tac toe game using your python.
We also saw that artifact will try to give you the interpretation of images.
Right. So I'm not going to do that check out. I try to interpret your charts and graphs
based on your artifact features.
Plot Sonic 3.5 is incredibly versatile and can be used in a variety of scenarios.
One of the standout feature of Cloud
Sonnet 3.5 is its natural language
processing capability.
You don't need to learn
any special commands or syntax, just type your instructions in plain
English and plot.
Sonnet understands
this makes coding accessible to everyone,
regardless of their level of expertise.
Imagine never getting stuck on syntax
this or. Having to. Write boilerplate code again
with Sonic 3.5.
Use cases of Claude
You simply describe what you want
to achieve and it writes code for you. Whether it's a simple function
or a complex algorithm,
cloud has got you covered. Debugging can be one of the most
frustrating parts of coding,
but with Cloud Sonnet 3.5,
it's like having an expert looking over your shoulders.

Describe the issue you're facing
and on it will help you and find and.
Fix the bugs. Quickly and efficiently. Plot Sonnet 3.5 Integrates
seamlessly with popular
development environments
like Visual Studio Code. This means you can use it
without disrupting your existing workflow.
It's all about making
your coding experience as smooth and efficient as possible,
whether you need to write a Hello World
program or a basic calculator,
class sonnet can do 10 seconds.
This is perfect for beginners
or for quickly prototyping ideas.
Tired of writing the same code
over and over? Plot sonnet can handle repetitive
task for you, freeing up your time
to focus on more complex,
interesting challenges. Learning a new programing language
or concept
class sonnet can help you understand new concepts and write sample codes.
It's like having a tutor that's available. 24 seven. Let's jump into the demonstration
using Triplet,
a powerful online coding platform
first with a triplet dot com and sign up for log in once
you are and create a new python ripple.
Give it a name like class on a demo
and click Create Ripple
by default Triplet will create
the mean dot by file for you. We will use this file to simulate
the capabilities of class sonnet.

So I am writing a prompt
in a simple English language that does write a function to calculate
and print the factorial of a number.
Demo on Claude Sonnet
I'm not going to the basics like give me a
hello world example and so and so forth. So I'm just trying to calculate
and print your factorial of a number.
So let's see what Claude
3.5 so on, it gives you.
So it's giving me the function. Right. Factorial is the function
if n is less than zero
return that factorial is not defined
for negative numbers. See so carefully
it has even dragged out your exception.
Like if a user is trying to go a number
which is less than zero, it is giving you or returning you
a statement calls factorial
is not defined
in the four negative numbers. Right? And then elseif end is equal to zero
and is equal to one.
You return one
and here it is giving you the function which is calculating
and printing your factorial number.
The number given is five. Then it will calculate
with the help of this function. And if we print that the factorial of five
is figured out, what's the answer
I need? A brief explanation is given on
how exactly the function works. Right. Okay.
So we will try another example. Like we also say that your class
on a 3.5
has debugging capabilities, right? So we will try to give an input

as an incorrect code
and we will ask Lord Sonnet
on how to debug this and why the code is incorrect
and give us an explanation on it.
So this is my prompt wherein I am having a function to add in B
and return E minus B lightweight
and I'm giving a function to add. But you can see that
I'm doing a subtraction over here, right?
So this should be A plus B, right? So I'm telling Claud that
let me know what is wrong
in this code and also provide
the corrected code for me. So let's see
whether it will debug this or not for us.
Well,
it says that I am happy to identify issues and give you out with the results.
So let's see what it gave us. So it says that
the problem in the code is an add
function, right? I just mentioned that instead of giving A-plus B in the code,
I mentioned that as a minus B, right.
So it is giving me the broken down like step by step
process of what went wrong.
The function is named N, but it is actually performing subtraction
and instead of addition, right.
It is giving you the complete details
of the debugging errors.
Great capability
you can try out for your complex tasks or even if you are a beginners,
you can try out with this.
Right? So it is giving me the correct code
and you can try that on your python.

That was about your coding capabilities. I have just tried out some simple prompts. You can just go ahead with your use cases

and try to explore

more on your cloud right now. We also said that your Class 3.5 also

has the capability

of logical reasoning, right? So we will try to see

some of the logical reasoning,

the prompts and see what

how the cloud will give us the answers. So this is my prompt

for logical reasoning.

It says that you are in a room

with three light switches, each controlling one of three light bulbs

in another room.

You can see the bulbs from the room

with the switches. You can flip the switches

as many times as you want,

but you can only go into the room

with the bulbs once, right? So how can you determine

which switch controls

which bulb are very good reasoning. Use case right. We will see what Claude provides.

Well, it does, given the complete explanation of

what is the reason behind

and how it is determining

that which switch is controlling which. Right. So let's see first thing like

it's giving you a step by step approach.

Let's consider what we can manipulate that does the switches

and what information we can gather. We know that light bulbs not only produce

light but also heat when they are.

It is giving you an information

about light as well. Right. Okay. Here is a strategy

to solve the puzzle.

Don't on switch one

and leave it on for several minutes. After this time, turn off,

switch one and immediately turn on.

Switch to the switch

three off then bedtime. Now enter the room with the light bulbs.

When you enter the room, you will be able to deduce

which switch controls which bulb.

Great explanation. Right. And then the solution works

because the bulb controlled by switch one

had time to warm up, but it is now off and you can see how exactly it is

giving you the logical reasoning

of the issue and it is giving you

the determine answer here.

Right. So we will try to do another logical

reasoning question, will try to give your and will try to see what the Claude 3.5 on

it gives you.

So I'm giving a prompt

which says I'm thinking of a number if you add seven to it, the result

is three times the original number.

What is the number? So let me think of a number first.

Let's see what Claude gives.

Let's call the number. We are looking for X Now let's translate

the given information into a mathematical.

If the see it says x +73 x. That's what the question said. Right. Three times the original number. Right.

So that's what it is doing here. Now Let's solve the equation. It is giving some equations.

I'm simplifying x just trying to give extra 3.5 and therefore the number

I'm thinking of is 3.5.

Great,

good logical reasoning at this giving. We also say that your thoughts on

it has automation capabilities, right?

So we'll just try to look a small example. We'll try to give a prompt earlier for automating your task

and we will see how it works.

This is my prompt here where

and it says that automate handling by writing a function to read from one file

and write its content to the another file.

So it's generating

it's handling errors as well. Okay, let's let's check out what it is

trying to give us here right. So the function is copy file

content, right?

It's like it takes two parameters,

source file and the destination fight. It uses a tri

x block to handle some errors and inside

try block does giving some

it opens the source file in a read more with the statement which ensures

that the file is properly closed.

You're reading obviously

would want the files to be closed. Right? And then we have the exception

handling cases

and that's how it is trying to copy

a few things to note. The function reads

the entire file into the memory at once.

For very large files, you might want to

modify it to read and write it in chunks. Obviously, if you have very large file,

you would want to modify that.

It writes in the chunks, right? It overwrites the destination file

if it is already exists

and you can just read out what it is

giving and you can as well explore more on this. Right.

So I'll just do one thing. I copied this code. You have a source text

and you have a destination D file.

These are the two external files which are being created
and with the help of your copy file content, it is trying to write
from your source, your destination file.
Right. So I'll just copy this function
on my on to my report editor and I will see what exactly
it is trying to give me.
So I'm trying to create a replica here
since it's a python code,
let me just write here as
automating some name here.
So I'm trying to create a player
to complete the project for me here. Now,
so you can see that
the project was created and may not be. Why is my mean Python file.
So I'm simply copying go code here. Right. Okay. This Is my code.
So let me just try to run the code so it is trying to give me an error
that one of the file that the source
statistics destination would see was not found obviously
because we have not created any file.
So we will just try to create a new file
called as source, not be extinct.
Okay. That's the file name. Right. So I will write
Hello from Cloud Sonnet 3.5.
So this is my file and I'm saving it. I'm creating
another file called as Destination,
the steam ignition dot, dot,
the destination or the dock. So as see that the file is empty.
So let me just run my file here so you can see that content
successfully copied from source
not to destination. So let us see if destination dirty
is it this copied?

Hello from Clerks on it 3.5. Now we are trying to have your performance

metrics

of your clock 3.5

sonnet comparison with the different

platforms

like you have your clock three Opus 84, Gemini 1.5 Pro and Lamar 400 beat

right So the graduate level

reasoning for zero short started part of this 59.4

higher than your Jeopardy four as well.

And then you have your undergraduate

level knowledge that is 88.7% for five shot and clock three has 86.8%.

Nice performance it is giving right? And then Gemini has 80 .5.9

and 86.1 for Lambda

that is less compared to your cloud 03.5

sonnet performance based on your undergraduate

level knowledge.

And then the code, right? It gives 92%. The performance matrix is 92%.

Then your class three three opus,

which is 84.9 and your good four

has 90.2 way

less than the clock, 3.5 Sonnet and then Gemini and Lemo

also are the back end like right?

Multilingual math 91.6 And comparatively

you can see that it is higher than your jeopardy

for and your Gemini.

Right. And then you have your reasoning over

text. Right. We did see some examples on it

and we did get good results on a try.

Try out few more reasonings on it. So it does 87.13 short chain of thought and then comparatively

it is less than the other platforms like GP photo and lama

and then mixed evaluation.

Again,

your class on it is on the higher end. That is 93.1%

and then 85.3% for your lama.

And then math problem solving and problem

solving capabilities as well. Right? 71.1% for your

math problem solving, which is comparatively less

than what your GP for all provides.

But still it is giving you mixed

good mixed evaluation performance. It is giving you code

evaluation performance. Right?

So still fair enough. And then grade school math. It is giving you 96.4%

so you can see that lot 3.5.

So note has the higher performance metrics

when compared to your cloud three Aupres de Photo and Gemini 1.5

Pro Mama 400 B.

Now I was talking about C or T, what

the C would be is your chain of thought.

So why not? Let's ask Cloud

eight on what exactly is your T? Right? So I'm just writing. Good.

I'll just give it to what to see what and

I will see what response to we get your

so let's see what it is giving it is giving me the like the definition of it.

So C or D typically stands for chain

of thought in the context of EML. It says that particularly in large

language models, it involves breaking down

complex reasoning tasks into a series

of intermediate steps or thoughts.

So we will see what are zero shots.

Zero short learning refers to a model capability to perform a task

or make predictions about class

as it has never seen during
training zero short learning. So it is giving you the definition of it.
So basically you can see that
the way you are giving prompts to your Claude I will help you in trying
to get a better outcome.
So this reminds me of one of the course on your prompt engineering
in our Great Learning Academy.
Go ahead to a great learning academy
and try to figure out what the court is all about. Try to give better prompts so that you can
utilize Claudia 3.5 sonnet as one.
Moving on to Claudia
Safety First, Claudia includes. Rigorous. Safety features to ensure responsible
high usage.
This includes
filtering out inappropriate or. Harmful. Content
and conducting regular safety tests.
For example, it ensures that a responses
to potentially harmful queries
such as how to rob a bank
for educational purposes are appropriately handled
to prevent misuse.
Claudia employs measures
such as data encryption, user privacy protection, bias detection,
AI mitigation and continuous monitoring
and updating of safety protocols
to ensure the air operates
ethically and securely. Today, we explored the features
and advantages of Claude,
including its butterfly capabilities and the innovative artifact features.
We also looked at how to enable
and use these features, as well as the robust safety

and privacy measures in place.
EIA assisted coding tools
like Cloud Sonnet can transform. Your coding.
Experience,
making it easier to generate code, debug functions and automatic tasks.
Give it a try and see how it
can improve your workflows.
GP Photo Mini Stand for Generative Pre-Trained Transformer
for optimizing one mini.
Think of it as the smart phone version
of a powerful computer
GPT 4o Mini and uses
compact, yet incredibly capable. It's not just a smaller version of GP
Foreo.
It's been meticulously
fine tuned to be super efficient and worse deal making it ideal
for a wide range of applications.
Imagine having an assistant
that fits in your pocket,
but can transform tasks
that used to require a full sized model
with the GP. Photo Mini you can draft emails,
generate creative content,
or even assist with complex
coding projects seamlessly. It's optimized for speed and efficiency,
meaning it can deliver powerful A.I. capabilities without needing
massive computational resources.
But the benefits don't stop here. GP Photo Mini supports real time
collaboration
allowing users to work together on projects no matter where they are.
It's at once. Natural language understanding

means it can help with research
by summarizing articles,
extracting key information,
and even providing insights
based on large datasets.
Whether you are a developer,
integrating A.I. into your apps, or a researcher
needing quick data analysis, or a student
working on assignment
or an AI enthusiast experimenting
with new ideas, GPT Photo
Mini brings the power of A.I.
to your fingertips. It's designed
to make cutting edge technology
more accessible and practical,
allowing you to leverage A.I. in ways that were previously unimaginable.
So why is everyone talking about? GP Photo Mini. This updated model offers
several advantages that make it stand out.
Why GPT 4o Mini
One compact and efficient. Unlike its bigger sibling GP photo,
the mini version is optimized
for faster performance and lower resource consumption,
making it perfect for devices
with limited computational power
to watch the title applications
from Blackboard to content generation
and beyond. GP for or mini is
What's the deal enough to handle a wide
range of tasks
without compromising on quality but cost effective with reduced resource
demand?

GP photo mini makes high quality more affordable, opening up possibilities
for smaller businesses
and individual developers. Next, we will explore a few features of GP
Photo Mini.
Now that we know what GP Default photo Mini is
and how it differs from GP photo,
let's highlight
some of its awesome features. Streamlined performance delivers
high quality
Features of GPT 4o Mini
outputs with optimized processing power, user friendly
integration, easy to integrate
into existing systems and workflows
scalable solutions
suitable for projects of all sizes from individual tasks
to large scale implementations
robust capabilities supports a wide array of functions,
including next generation
language translation
and, award financed accessibility
available to a broader audience
thanks to its efficiency and lower cost.
Real time ideal for check boards and other interactive
applications needing quick responses,
customizable framework easily tailored to fit specific use cases and preferences.
Now let's look at the key differences
between GP default. GP default or mini.
Difference between GPT 4o and GPT 4o Mini
Size and speed. GP Default or
mini is designed to be smaller and faster.
It processes this information

more quickly, making it ideal for real time
applications.
Resource efficiency. While GP 84 requires
substantial computational resources,
GP Photo Mini is optimized to run
on less powerful hardware without significant loss in performance
accessibility with GP photo Mini. More users can leverage the power of GP D
for technology.
It's accessible on a wider
range of devices from desktops to mobile customization.
GP photo Mini allows for easier
customization and integration
into specific workflows,
making it adaptable to various needs.
Ready to explore GP Photo Mini. Head over to the official GP Photo
mini website.
Click on Get started
and you will be on your way to integrating this amazing tool
into your projects.
GP A4 and its variants, including GP Photo Mini,
are generally not available for free use.
Opening Eye offers access to its model
through subscription plans or API usage,
which typically come with a cost.
Some limited free access
might be available through certain platforms or promotions,
but sustained and full
featured access usually requires payment. Now we will explore the open air platform.
First, let's explore the playground. This is the main area
where you can interact with and test
Demo on Playground tab, Dashboard tab, Docs tab and API References tab

different GP models. You can input, text
and see how the model responds.
Next is the check this section,
especially for stepping up the testing conversational models
like Shadow GP.
You can provide system instructions
and see how the chat
model responds in a conversational format. The drop down GP three point fight.
That book lets you choose
the specific Watson of the GP model.
You want to use assistance. This allows you to create and manage
different
assistant personas or configurations. You can define
how the assistant should behave,
what should know,
and it's specific purpose. Text to speech.
This feature converts written text into spoke in words.
It's useful for creating wise responses or annotations completions.
Here you can generate text completions
based on a given prompt.
This is useful for tasks
like content generation video provide a starting sentence
and the model completes a forum. This likely refers to a community forum
that users can discuss
various topics related to open eyes
models, share insights, ask questions and get support from both the community
and open air staff.
Forums are valuable resources for learning best practices, troubleshooting issues
and connecting with other users who are working on similar projects.
Next, we will explore the dashboard. Fine tuning. Fine tuning allows you
to customize the GP model

by blaming it on your specific dataset. This makes the model more accurate
and tailored to your unique requirements
batches.
This option lets you process multiple prompts at once.
It's useful for handling large or running multiple tests simultaneously.
Storage. Here you can manage your saved data
prompts and responses.
It helps keep your work
organized and accessible.
Usage. This section tracks
how much of the service you have used,
such as the number of requests made or the amount of data processed.
It's important
for managing your subscription and staying within usage limits.
API Keys. API keys are used to authenticate
your applications
with the open EIA API. They are low. Your software securely access
and use the GPP models.
Next, let us explore the dock step. So here you the exact model names
along with this description.
So you have Chad GP four along with Chad
to pretty photo mini meetings.
You can just go through the model
and its description with its context window
and what are the output tokens
or the training data
required for port demands. Not just that you have the descriptions
of other models
as well, like GP, T4, Turbo and three point Turbo and many more.
You can just go through all of it. Next, let us explore the API reference
step.

Here we have the introduction of API through HTP requests from any language.
So how to install, how to give the
commands, how to give the library commands
and what are the authentication
API keys and many more. So you can just
before these things as well.
These will be helpful in giving the API
requests in your prompts.
So now let us go back to the playground. Try giving the
Prompting on Playground and Billing Settings
prompt in the completions section. So here, if I ask any question
or if I write or type anything,
it should actually continue or complete
the sentence and give you the results.
So let me try doing it
in the free account. So we'll see
what is possible by. So when
I type something and click on Submit,
if you have the paid Watson,
you will be getting the result. Since I am using the free wasn't here,
it is telling me
that I have reached the usage limit so it is directing me to the billing
settings.
So here you can see
the billing section of your account. So for free trial, I have zero credits,
which means it's free.
It's a free account. That is why
I'm not able to use this platform. So you have. Your payment.
Details
where you can go and do the payment. It is some dollars to use your GP for
and for me.

So these are the things. So here you go. You can do the payment
and get your subscription and try to use
GP for mini and explore
more features by yourself. In a world
where technology is constantly evolving,
GP Photo Mini stands out as a game changer, making advanced
AI more accessible
and practical than ever before. Whether you are drafting emails
collaborating on projects
or diving into complex research, GP Photo Mini is your versatile
efficient and powerful AI assistant. It's like having the
brain power of a supercomputer right your pocket ready to tackle any task
you throw at it.
Embrace the future of air with the GP. Photo
Mini and unlock endless possibilities
in your everyday life.
Version of Google Gemini
In Gemini, we have different versions
broadly classified into three elements.
The first one is Flash, then by Nano
and then Ultra or Broke.
Let's see,
what are these classification mean to us? First, when it comes to flash,
it was developed
in order to execute the queries
or the branch in a very quick orbit.
If you got a nano right, this a variation in order to execute order, use Google
Gemini in smaller devices like Fords or Tabs Nano motion was discovered.
Then we have our grant
and when you come to this version, it is having capability
to execute biggest plants

and also give you a good detailed analysis
according to your requirement.
Compared to nano and flash generally,
you could. Nano has a very limited number of tokens
compared to Flash and Ultra.
So this is the variation in broad
you can have in Google, Gemini.
Then let's explore versions one by one. You can see
Gemini started with version one IGI,
but they launched a product called BART in September 2023.
What are the tokens of 1 million? It was generally used in order
to have initial conversation,
a model and also next generation. It was specifically given by the Google
as a product,
as a competition to open air. This job
depicted the applications of Gemini.
One is just to have a conversation
with the chat and also basic
content generation. Then it also served a general purpose
II pass as well.
After Gemini one that is Bug, we got Gemini advanced.
Again it is BART plus it was in late 2023 after September.
It was also having tokens of 1 million, an enhanced performance
compared to Gemini one.
And also more refined responses for
this was available. Improvised chat bots and more complex
content creation
could be handled in Germany. Advanced. That is BART plus,
after all these, as I mentioned, we got the broad classification
that does nano pro and then Hydra. So let's understand
what is Gemini Nano 1.0.

So it was in the early 2024
it was released. But the tokens, if you could see it
as half million, meaning
it is having a less number of tokens comparatively to other versions.
Why the reason? Because it is
generally used as a light weight model.
It is on the devices like mobiles or Iot devices you can't execute or use.
Gemini Nano 1.0. Generally the applications are
if you could see
it's in mobile
apps Iot devices and embedded. It was specifically designed in order
to work as a lightweight model. Then comes Gemini Pro 1.0 again
in early 2024
having 2 million tokens. Was used for professional grade
of artificial intelligence
and also multi-modal capabilities in general.
This application was only preferred
to watch professional content creation
and advanced research.
So this is the application of Gemini 1.0.
Now let's go to the next one. Gemini I 1.0 early in 2024.
It was also having 2 million tokens
high performance model,
and it was used only for intensive tasks.
It was generally it was for intensive
tasks, complex problem
solving or high demanding applications.
What the use of Gemini Ultra 1.0 Gemini 1.5 Pro was again
early version of 2024. It is also having 2 million tokens
generally used for advanced multi-modal
reasoning and also long context. Window

air driven content creation
is the application with our data analysis feature and advanced coding features
available in 1.5, which is currently.
Most. Used and. We are also learning
more about 1.5 pro in this video.
Stay tuned for that. Then we have Gemini 1.5 plus,
so it's 2 million tokens. Again, Max, it flash means, as I told you,
it is having a speed,
it is having low latency processing
and also is generally used for mobile applications,
a smaller device.
It's not restricted only to mobile,
but for smaller devices. Then it is
generally used as a realtime application.
And IAP integrations were done
with the help of Gemini, 1.5 flash and summarization
and extractions.
All these other applications of the Gemini
1.5 flash. Hope we are clear till now,
but the variants of Gemini.
Now let's see what is the future? What is Google preparing in order
to get the next version?
That is Gemini two and 2.5
and we don't know much of the information how when it is released for the users
and how many tokens it would be having,
what would be the key features
it does just like advanced version of 1.5 flash advanced, multi-modal processing,
higher accuracy
end image and we do understanding
cross domain AI interaction. So all these would be the features of the
and it is generally used for video

and multimedia editing tools. If it comes actually editing
would be very easy comparatively.
And when you come to 2.5,
what happens here? You have advanced
predictive analytics skills
and you can do a lot of data
processing, data analysis, how multilingual communication
support as well.
So customer support would be very easy
if this 2.5 comes up
because it is having
a multilingual service as well. Now let's switch
back to our Google studio.
General Prompt Demo Google AI Studio
Now, let's start with that first prompt. You have to click on the create new prompt
and if you want to change the name,
you can just go on edit C demo one and click on See.
Now the interface is
ready to take your prompts, but there is one option here
System instruction why this is used. This is just to set
the tone of the prompt, the responses
we get and the style of instructions
we give to the model. We are just trying to set a context
to before we stop
communication with the model. I'm trying to tell the system the response
we receive should be
for a right. So it is just a sample.
I just give HI as the prompt and run.
Once I do this, it has taken two tokens in order to work
just to see a reply for height. It is giving enough manner right?
It is giving a complete of the super soft greeting brief. Hi.

You see a single leave so it is trying

to give a poetic response to it.

So if I see how are you then it will try to give us answer enough

for the matter itself.

You could see the token count here. It's moving on and on. The time it processes the number of tokens

it will use, those are kind of building

blocks in order to process the statements

we give to the system. So you can see for just one, how are you?

It is giving us four lines. Answer because I have set the context. That answer should be in poetic manner.

The theme should be put. So this is o gender prompting,

which you can do

setting this system instruction

with Google AI Studio. Now let's move on to new build model.

What is this all about? Before that, see it does giving us Do you want to see this?

Yes. If you want to click on okay, it will save us demo one

Once you select the Newton model,

this is where you land. You have the option to create a structure,

prompt

Structured Prompt in Google AI Studio

the prompt

which we gave was a normal prompt. You have a second form of prompt

that is structured prompt.

So I click on structure prompt again,

if you want to change the name for this,

you can really put the name I'm using the name demo to I'll save it.

You could see there are two sections here, but section is input,

then output label you have textual prompt.

You have to first read the system,

then you have to best the system. This is the concept

you are doing in a structured manner.

You're not just giving a random prompt. Come on now let's set a context
or style of instruction.
The first thing the structure prompting. Let's set an optional tone
or system information with
you want to save it across? I'm getting the overall context. I'm working as a
news reporter. I need the normal news
to be displayed more excited
manner or as a breaking news.
So when I do this, I'm setting up context. I'm just saving. For now,
it has been saved.
Now let's start giving that input
and specified output for the model business,
trying to paint the models. So.
This time Olympics is held
in Paris right?
Another import. The player from
Turkey won a silver medal in shooting.
So I give this next
it will try to suggest to certain things.
It will just leave as a disk. Now let's type the response as the
did you know the Olympics
is held in part. It's exciting, right?
So this is a normal sentences
which I'm using. Then
want a player from Doki won
silver medal in shooting.
Isn't it exciting? So I've just typed
certain inputs and outputs
that is training a model out
to find examples of use to sample.
And you can give a third one as the

India lost medals in badminton,

right? You can just give off simple input and

oh my God India missed chance

to bag medals in badminton.

It is just exaggerating or giving an excited output for us.

Normal sentences. This is how I am trying to say. Then we request the inputs.

I see India won gold in wrestling.

Now let's try to run to get the output I'm clicking on run.

It gives exaggeration here in the outtakes

Gold in wrestling. This is huge. What a performance.

We are trying to exaggerate. First, we try to train this model and then

it is now trained across. It is also giving. As we mentioned,

we need the output an excited manner.

So this is the structure prompting. Now what do we do? We try to build this right.

The outputs will try to delete and we will just give the same input.

India won gold in wrestling and run this hundred percent.

It will follow the system instruction

which we have given at the top, not the model, because we don't have

any samples here in order to compare.

So it is a big paragraph. Breaking news India claims

gold in wrestling,

so it just try to give a good information. Often you so this is how destructive

prompting the book.

So now we have understood the normal prompt

and the structured prompt.

Let's click on new the old model. Before that,

if you want to save, you can click on okay

Model Tuning in Google AI Studio using System Sample

and you can produce the model. You can import that it should be RC

As we find on a Google sheet file

and we generally recommend how to find it.

Examples in the cheat available. If you want the modeling tool guide,
you can have a look with this document.
Now if you click on the select
data source, you have structure prompt,
which is the motto
which I created recently that has also been available here. And you have news headlines
and summarization, which is the samples.
So I prefer taking news headlines. Once you do that,
you can see the prefix column available,
you can select the prefix column,
input and output. This is how it will be
and just it is a sample.
And then you can change sample
to your model. Once you do this,
you can give a description if you want to
and we are not going to touch this. Advanced settings.
How many tuning of a box should be? What is the learning rate
multiplier and batch size and stuff? Let it be as it has been set
and you have two options here.
That is one point or pro not not one and 1.5 plus not not one tuning.
So generally we were walking
with the help of probe. So I'm using pro again, if you want to be,
you can click on the unit button.
Once you click on the tune button,
you could find sample model available in your library.
Once you click on the sample, you model
towards your left, you can see the
different model results. It takes some time in order to load this
and you have to wait for 5
to 6 minutes. Why? The reason being I selected it took
certain time in order to load this box.

See, the graph is going in the right way,
but the value is 1.2. Actually,
it's supposed to be at least 4 to 5.
That is what we consider a good value. But this is how model tuning looks like
in Germany.
As studio, you can also see less than one minute left
in order to finish this process
and it has all the information available
the box, the epochs, the learning rate, the examples
it has been doing for the graph
as the last graph
has also been put up here. This is the overall view of model tuning
which you can do in Google studio. Coming back to the main interface,
Other data import options in Google AI Studio using System Sample
you have an option here,
which is plus if you click on plus you have various available
plenty of options in order to choose.
But that is my right upload to record
audio sample. You may just end sample video.
So let's
check out the help of sample video. There are predefined videos here.
I'm trying to select one of the video
and add a prompt. What is this video all about?
So I click on Run. Once we do this, it will try to analyze the multi media
that is the video and give us the details.
What is this ten of dealing with? It takes certain time. If you could see the tokens of this
taking in order to process.
This is huge
because of the 10 minutes of video. So the responses in those lines,
it is starting to give a gist of
what does this video all about. So you can also upload the videos

from your local system
and understand
what is happening in that video. It has the capability to analyze
the video as well.
And you can change the model here, order to immediately get the responses.
See Sample video. I just gave the same video or another
10 minutes video.
How does this particular viral load and how fast it will work?
What is in the video exactly? I'm trying to give this option.
I'm clicking on run. I'm just changing the model and checking
how does this work? Why we were executing the second video.
We turn this model into 1.5 plus. Let's see what is the difference here?
In the first video,
it took around 30 seconds in order to work here we got 40 seconds.
And it depends upon
what is the content with you? What is the file size of the video?
Everything. The more content you have. It's good to use flash, right?
So it just a simple comparison
between pro and flash. You can do the same thing
for sample images.
If you click on
it means you have total here. So if I ask what is this,
it will try to give us the information let's see how much time this will take
when it doesn't flash.
It took around 20 seconds in order to say
this is a picture of us.
So this is how you can lawyer
images, your videos, your multimedia and understand
you have to first upload to the drive
or you directly take the information
from that. Right? So this is how the prompting looks in

Google studio.

So you have to choose it

from the available options that is from your drive

or the sample videos on record.

The audio as of. This,

let's get started with our demonstration. We will be building an Android application

that generates

Abstract of the Email Generator App

email for sick

leave and casual leave using GP2. As you know that we are going to create

an email generator app.

First,

let me tell you what are the different softwares are required for having

these are being generated.

Since we are implementing an Android app,

we need to have an Android studio editor.

So for that

you need to go to the website of Android studio editor which is developer

dot Android dot com studio.

Software Requirements for App

And then when you open the website

then you can see a link called us download Android Studio.

And once you click on download

it will ask for your terms and conditions to be accepted.

So you can read these terms and conditions

and in the end you can just click on. I have agreed on the terms and conditions

and click on download

and you can see that

the file is being downloaded. Once you download the file

from your Android studio website

that you need to see this Android

Studio 2024 and the version of that. So you have to click on this
and execute the steps which it follows.
Since I have already installed
the Android studio. I am not going to follow the steps,
so you can just carefully follow the steps
which it is giving on the screen and try
to install the software ID on your laptop.
The next thing which we will be doing is
we will try to create an epic key here. So to have this epic, I'm using Chad
Djibouti
open the AI platform
for creating this key. So I have already logged in to my Jack
Jupyter Open API key.
Even if you have a free version
or a paid version, it's fine. Version will have the usage limits
for your API keys, but still
you can be able to use those API keys
and have your apps being generated there. So as you open platform doors,
open the dot com.
This is the page. What you see
when when the website is open, you have to go to your log
in to your user profile over here.
And then if you see over here
you can see user API keys, right? Since we are trying to create a new API
key, we need to set create new API key.
This is optional, so I will
just give it as app here just to say that I'm creating an API for this
email template generator app here.
And then when I click on Create Security,
the key is being generated. So we have given the permission
to read and write API resources, right?

So this secret API key is needs to be kept
very confidential only because there can be some data breaches
which, which can happen.
So you just have to copy this
and keep it very safely. So I'm just trying to open a notepad over
here.
I will try to have this generated key
saved over here so that I can use that in my program. But I'm trying to implement this image generator app.
As you can see that
my Android studio is being installed and we are opening
the Android studio here.
Implementation of the App
So now we will try to create a project for your immune generated
app using your elements.
So what is the first step that you need
to do is you need to go to file, go to new and the new project,
and then you can just click on empty
Activity and go to next over here
and then you can give a name for your app. You're I'm trying to give or a name like Email Generator app. Okay. And you can see that the package name is
called or exemplary means Innovative app.
And this is my save location and my SDK is the default SDK,
which I'm using over here.
And my build configuration
is called Calling the Ascent, which is the recommended version
for this app.
I'm not changing any of the default configurations here,
so I'm just trying to click on finish. This will take some time
and a new project will be created.
So now you can see that

the project is open and we have two folders called us App
and then great script.
So we will see one by one
what are these folders and? Help us to help us in trying
to implement this app with you.
So we'll just open
Gradle scripts over here. We have some of the configuration files
you can see over here, like Built
or Gradle ideas and then you have lives
dot version to deployment.
All these files are actually required
for having your configurations being run. Okay.
The only thing what we will be
changing over here is your build or gradle is fight.
This build out Gradle file will help you
in trying to add those configuration design the required dependencies
which will help you in trying
to have this app being generated. All right. And then we will also
this build out gradle file
will also help you in trying to figure it out.
Or you can say debug
or try to test your tool and help you analyze
with how the configurations are
being saved and how the dependencies
are being used with it. Right? So the only thing which we would be
changing over here is we will be trying
to add the dependencies
which are required for our app here. Okay.
So you can see that I have just added
the dependencies of your android
and you need an express. So all these dependencies

I have added based on our app here

and once you have modified your grading file,

you need to actually sync your project,

which will help you in trying

to work properly with those dependencies. So I'm trying to sync the project now,

so this will also take a while,

but you can see

the sync is successful or not. If the sink is successful, it means that

the dependencies are added successfully.

If not,

you need to check out on the arrows and try to debug

what sort of errors of it. And you need to figure out

what the problems are and try to fix them.

Basically,

you will not have any errors there. So because

we are only adding the dependencies

now, you can see back the sync

successfully. If the sync is not successful,

it can show you errors in this

problem's dominant one here. There are some warnings, but we can just leave those warnings as such over here

because it's just setting us

to update the new version. But we are just trying to use the versions

which are apt for us.

Image and Beta App now. So that's all about your written script. Now moving to this folder app,

which is our main folder,

which will help you in trying to add

those functionalities to your app and you will be able to add those

of input, text screens, your buttons,

all those things and how to interface

between your front and back on all your

with the help of your API keys.
All right. So you can see that in app
we have manifest file, we have gotten double file
and then we have your rest file over here.
So will just explore one by one
what exactly these are and what are the changes. Are we going to have been incorporated
in these fights?
So in your manifest file
you have manifest and some insight. So if you see this Android
manifest or symbol file, this file will help you
in trying to give the critical information
for your Android system.
And also it will help you in trying to access the permissions
which you're trying to give over here. So we will just try to of not modify
much over here,
but we will just try to add
the connectivity with the help of one line that is your Internet,
as we try to add that over here.
So you can see that the only thing which I added over here
is the user permission button.
I'm telling that the user's android
needs the permissions since we are trying to interact
with the API as well.
So this is the only one line change
which we are doing in this site. Okay, next we will move on to the Scotland
or Java file.
This is our main package that is your com direct
example dot email generator app.
So we'll just try to explore this. What exactly does that this is our main
activity dot collecting site.
It is actually the main entry point

for your app and you will be trying
to define the user interfaces. You will be also giving the API give it.
You have just generated and tried to have
that functionality being of this, right? So we will first try
to create the user interface,
then we will try to add
those functionalities of your raw fetching API is in to have those requested
responses in your note keeping.
Right. So what exactly do we have to do when we are trying
to have the layout set for your app?
So you just have to go to this ideas folder over here you can see that
very few files over your enjoyable map
app values. And so we will just try to create
a folder called as Layered Folder.
So when you're trying to create just right
click on the rest folder and then go to your Android resource
that actually and then you just try
to give your resource type as build
and then click on. Okay. So, so you can see that
the folder has been created over here.
Now you just have to create file
over here again. We will right click on it, go to new,
go to your resource file over here.
It will ask you for the beam. So we will just give the name
as activity main
and then you can see that
the time has been created. So this will give you the layout
of how the app will look like.
So I'm just going to go in the constraint
layout and go to X and then because I need to do some modifications
in trying to add your X buttons

and your review buttons over here. Right. So we will try to add those text
and redo buttons over here.
Okay. So you can see that
I have created a linear layout over here and I have an edit text
and then I have two buttons over here,
one for generating the sick leave email
and the other one for generating the casual leave email.
And then I have a text fuel
which will actually give the response in this text view. It will give the response of my generic
casual be me mango security
and what's the next thing
which need which we need to do. The next thing which we need to do
is we need to try to modify
your colors strings. I'm your theme starting seven five. Next thing we need to do is we need to
check out on the user interactions.
We will try to modify your colors,
simple string, dotted symbol, and then also try to add
some of the styles for your experience.
So I'm just opening colors dot symbol,
and I'm just trying to find it based on the details
of your colorful minds, based on how
I want that could be as this string board
simply manages your X resources,
the color not accessible manages
your color resources. Now we will try to create the style
X symbol again, right click on New Values
Resource file over here and then you can
give the finding as states.
Click on Talk of Files opened All this. This file will help you
in trying to manage the team.
Your text. Now that we have created the layout,

we will just try to add the API
which we created in our main activity file
so that we can have that interface. We generate.
So this is our main activity defined here. You can see that
I created a private function over here
to send the query and I'm using the
you are unless you could be as and when you babble.com which is the
you are in which we are using your DPA
in the model which we are using here
is you could use 3.5 turbo and then the prompt is equity, and
then the maximum token would be defined
as is one to be over here. Right? And then we have defined the request,
request body
and all response
words, body has to be given all here. So it's just a simple file read and
you are just trying to use or reproduce.
Try to send a request
and try to retrieve all response by that. Right. And you can see that you have created
a variable called as the API key
and the generated API key,
which we had generated using our charge. If you do, you can just copy paste
the API here right now.
We will try to execute the finds and we will try to
see if that was generated
on the Android mobile or not. Right. So to do that, we will first try to
figure out if there are any errors or not.
So we will go to the file,
we will go to break. We will first try to clean the project.
And you think sometime
I'll give you a clean the project now. Now we will try to rebuild project.
So now you can see that

the Bredell bill has been finished. So you can just see over here

if there are any errors or not,

you'll be

we will figure it out here, right. To run this application,

we will be using an Android mobile.

So I will be telling you what are the settings

which you'll need to do so that you will be able to access

your Android mobile for this app.

Let me tell you how the settings

you need to do when you're trying

to have your Android app

Executing the App

as your

physical device to run this up, right? You just have to go to settings over here

and then go to your about phone. And after you have clicked, you're

about to go to your software information.

And here you can see

there is something called as build number. So you try to tap that build number

and it will tell you

how many times it has to tap

for the developer options to open. And then it is asking for my finger.

I'm just trying to give my password here of the mobile app.

And then you can see that

a message problem like developer options have been enable.

So how do we check that? Go back to the screen and then you can see that after you're about home page,

you have your developer options.

We can start here, right? So just click on your developer

options over here. And then after that you can see

that the developer options

have been on over here since we need to sync your app,

whatever you have constructed in your
Android ID to this mobile. So what you just need to do is click on
any will You always be debugging
you need to enable
your new is pretty buggy so it will try ask you whether
or you is be debugging reasons
intended for development. You just say okay and then no will be able
to access your device
through your Android device. So I'm trying to click on. Okay.
And now we will click on the run here. So we see that the project is big and
it is launched on for your Android mobile.
Now you can see that
the app is successfully installed on your mobile.
So we'll just check this out
so you can see over here. So I'll just try to give a prompt here
that I have before
and need to this the rest. So after this
I'm just trying to click on Generate
and when I click on Generate Catalina, you
maybe can create the guys in this room. So this is how it is
easy to build your Element
app on a simple app
is what we have created here. We have covered large language models and
explored some of the most popular ones.
That is deeply flawed and Yemenite. To learn more about islands,
be sure to subscribe
to our channel
for exciting videos and updates. Next up, we'll take a closer
look at some of the most popular
identity tools
that are Transform industries. Are you ready to discover

this game changers?

Let's get started.

Generative A.I. is here to spark your creativity

and double charge your projects,

Generative AI Popular Tools

whether You are an artist, a writer, or just someone looking to jazz up

your content.

These tools are like your new BFFs

in the creative world. Now let's talk about tools.

These are gadgets that make generative. I feel like magic and we will be exploring

this in this video.

First is Chad, Djibouti. Chad, Djibouti was developed by open air.

ChatGPT

It is a state of the art language

model renowned for its conversational abilities

and versatile applications.

Second guitar copilot powered by open

air codecs.

Github Copilot

U2 Copilot is an able programmer

that assists developers

by generating good suggestions

in real time next year.

Claude

An assistant specialized in ramp

engineering

FLOAT offers advanced capabilities for generating text based on user prompts.

Gemini

Last but not the least, Google bot, now known as Gemini, is Google's

innovative A.I.

platform,

which seamlessly navigates to text gold audio image video with capabilities
like multi modality.
Throughout this video
we will dive into each of these tools
uncovering their functionalities,
applications and the limitless possibilities
they offer in the world of generative
knowledge to understand what is prompt engineering.
Prompt engineering is as much an art as it is a science.
Basics of Prompt Engineering
Oh, very creative line. We'll come back to this once we understand
what is actually prompt engineering. Right. Engineering is made up of two words
product and engineering
brand is nothing but a detailed set of guidelines or instructions
given to the LLN
or generative model to do a task. Engineering is developing iteratively
a task specific product to enable the generative models or elements
to output a perfect or new perfect
outcome that you have focused for. So if you see brand engineering
is an iterative process, how first is the idea that comes to you
that you want this?
Then you have a
prompt, you design a prompt,
then there are results. You give that to your model,
your child, you look
and then you have the test,
you have a feedback, right?
So basically it's a process. Iterate process.
Now once you have the prompt,
then you have the results. You might be not very much satisfied
with the result.

This is where you develop or you have a conversation
with your child.
You party with your model and and force or reinforce
the model to give you a good result, a good output, which is what you want.
For example, if you have a task on gold generation now
a specific problem statement can be solved
through various methods.
For example,
you want to output palindrome.
Palindrome can be done by brute force. Palindrome
can also be done by a recursive method
recursion. For example, when you give chargeability
a prompt, it should first
give you a brute force method. You have to keep on instructing
and providing an
attempt
or providing a prompt to charge you pretty so that it gives you
an optimized version for your code.
This is an example for code generation. Similarly, suppose
you have content creation.
You want to create a content,
and your content that you want to create is directed
towards
a specific target audience. Now then you give Chargeability a prompt.
The results will show you
that might not be satisfying enough. You have to tell your child
GPP audio generate is more iterative. Be what you want, give them a feedback.
For example,
you have not received a good output. What you would tell your
chargeability for example,
this is not I want. Please generate

the output based on my target audience or based on my feedback provided.
Right. That's where your feedback
or test comes into picture. You have to have a conversation
with your generative model
or in case
if you see Jeopardy, give them a feedback. Constant feedback.
It is good, it is bad,
It is what specific part you want it to be changed, right?
All of these things
are an iterative process. That's why we see
brand engineering is an iterative process
based on the guidelines
or the detailed set of instructions that you provide for a specific task.
Now let's go back a bit. We said when we started
about prompt engineering
that it is an art as it is a science. How it is an art.
It is an art because when you conceive an idea
and when you write a prompt,
it is your creativity
which comes into picture. However, the science part comes in
when you provide
your generative model or element, or in this case, tragedy, the prompt
and it gives you the result.
Because in the back end
there is something called generative models,
there is something called different
transformers or architectures
that are going in the picture. Their billions of parameters
are playing a role
to give you this result. That becomes your science part of it.
That's

why prompt engineering is combination of both an art and science.
Now that you have understood
what is prompt engineering, let's go and see
the vital
the main element of prompt engineering that is from without which generative
model charge you but you will not book
prompts are constituted or it contains stubborn force
majorly that is perimeters and structure. What will be the structure
and what will be
the parameters of your plants
that when you design it, you have to think so
that you have an optimized result.
Now, structure is one thing we'll talk more about and parameters here.
Let me explain. I have only selected few parameters. You can search more.
There are more parameters on which
want work majorly.
Here we are highlighting
only three parameters that is temperature. Don't be and max length temperature is
the randomness that comes into picture. Then you provide that parameter
to your prompt.
It ranges from 0 to 1. And if you want your model to be creative
and generate creative answers or generate creative outcomes, your range
or you set your temperature
to 0.7 to 0 point. However, in tasks such as gold generation,
you don't need creativity,
so you can set your temperature to zero. Also there you do not want creativity.
Now don't be is actually kind of seen as empirical.
It also has creativity. Dopey is top probability
that it selects from.
So you have a generative model. It generates results based on your prompt.
Now there can be any number of results. How does that model decide what

to put as output for you? You stop becomes into picture

more the copy more you have different creative answers, right?

So low for factual, high for die was last is max length.

Again it is written. Manage your response length. This parameter is to control

the cost of your generative model as well.

So now you have understood

what is prompt engineering,

what is a prompt and what are the two main components of your prompt

that despite immediate and structure.

Now let's see. Components of a good prompt. What we mean by good prompt

is that your charge,

your generative model

gives you a good output, a good outcome,

and you have less feedbacks for your child GPT or generative model.

Now let's see the first or I would say the first are context

and instruction.

These are interchangeable. How do you want to design your plant? If you want, you keep instruction first

and then context in your prompt,

you can do the same. If you want to give the context first

and then instruction, you can do the same.

Context is an additional information

that you want to provide to your model.

Instruction is a specific task that you want your model to perform.

For example,

if you want to summarize some text that is your instruction, summarize

your text, summarize This text.

Context is why you want to summarize what is the outcome you want to achieve.

For example, going back to Tesla, there

you are

providing an instruction to summarize, but the context is that you want all

business report from a business report

you want. What is the profit from 2021 2023, right?

So that is the context you need to set. Why do you want this? Why it is needed.

Next is the input data. Okay. What is your input?

For the example, this example,

the input data is your business report. A business article

that you will provide to target

and output indicator is what, how, what and how you want as an output.

For example, you want in a CSP format

your output the boss you want in a tabular format

or you want in a graphical format.

So you have to tell what kind of output or what format of output you want.

So these are the four components

of a good prompt. If your prompt consists of these four

things,

you are enabling your generative model, your chargeability

to give you better and good results.

Now let's see an example here. We will see sentiment analysis.

We we are setting up a context. It is written act as analyst.

You are acting as you are asking

your child you to act like analyst working for an audit platform.

You are seeing your child

that you even have to perform sentiment analysis based on feedback

provided by the consumers.

You are working for an audit platform. Your consumers are the users that are

seeing your movies, your VP series.

Now, the feedback that they provide on

that, you have to do sentiment analysis.

This is the that you are giving

to your child, depicting not giving this.

Your child typically knows that they need to think

from an analyst perspective.
They need to do sentiment analysis. Now, second is instruction you are saying to Chad,
you need to classify the feedback
into neutral,
negative, positive. Then you are instructing
or telling your child.
You pity that what is positive. Positive means
your consumer is your promoter.
Negative means they are the model. Neutral means
they are neither promoter or promoter.
They are neutral. Your content. This is the instruction that you give
what the charge you,
what generative model needs to do. This is what it is doing.
Classifying next is your input data
and your output indicator.
Before that, let's see an example, a good prompt
video giving a good prompt. If you give examples for your generative model
your In this case
we are using chargeability. If you give them an example,
they will learn from this example and clean your input
data to give the output. Here we have given two examples.
Feedback was feedback as I think the CDS was okay.
Second feedback is the acting of each character in the series was awesome.
So I think the CDs was okay. It took this. Okay as new.
So this is a sentiment neutral here. Awesome.
It took as positive. So your sentiment becomes positive.
Now your input data is the story line for the series was repetitive and abysmal,
and your output indicator
here is sentiment. What is the sentiment for this?
You're asking your Djibouti. What will be the sentiment
of this entire project?

That is your context that you have given
which discharge you did to act as an analyst
and do sentiment analysis.
It gives an instruction to classify the provided
that is your input data and classify that feedback into positive
or negative and neutral. Here you and you
you are giving examples for the scene
so that your Chad Djibouti can understand, learn and then act accordingly.
Now, the output for this particular feedback is negative,
so if you copy paste this
into your chargeability, it gives you output as negative
for this sentence. But this is how
you should write your prompt,
which is good. These are the components, good components,
all your prompt
so that the output that chargeability generates for you is
exact for you to work upon. So this is what we understood
about components of the good prompt.
Now that you have seen, what are the
different components of a good prompt? Let's have a look at a good checklist
that you can keep in mind while designing
effective prompt, but better results. So the first one is defining the good
billing. Chad Djibouti. What exactly you want to do that is going defining the goal
next is in detail out the format format is output format that you want
Chad Djibouti to provide you. For example, as I told you previously,
it can be table paragraph, list us
vs anything that you want
and you can give it in a priority order.
Also, ask Chad Djibouti
to give you in a priority order if you are working on such content.
Next is creating a rule. If you remember the previous prompt

that we discussed
while learning
about components of a good prompt. The tool as a context act
like at analyst that is creating a role so that
Chad Djibouti assigns our processes based on your request,
based on that role
that you are asked. Next is clarity who the audience is
or clarify who the audience is. Basically, you are telling to deputy
leads to the output
based on the demography of my audience
who I am catering to.
It can be beginner level, intermediate advanced
if you are creating any content.
If you are creating code,
it can be foundational functional export
based on the audience. It can also, if you're a teacher
and you are using chargeability,
you can specify. Why I want to teach fractions from oh
six plus learner perspective, or if you want to create a content,
or if you are working on something
and you want to make someone understand
that particular concept.
For example, I want a ten year old to understand what is prompt.
How would I make that particular learner understand
that kind of clarity,
that kind of clarification? Then you put in your prompt,
what is your target?
It will generate tailored feedback style output
for your prompt. Next is giving context.
Basically, you are giving every possible additional information
for chargeability in your brand

so that the purpose of your request
is clarified to charge you. And the response that it generates is
what you want. Next is giving examples
we saw in the previous components
why we would understanding components
that giving example
trains your charge you makes it understand that what you want
and it learns from it to produce
accurate results. Next is specifying the style. It can be communication style.
If you are working with a brand,
it can be how your brand works. Communication style
such as informal, formal.
What do you want to do? So it's suitable for your response.
Next is defining the scope. Then you outline the scope
with specifications
besides giving context and examples.
Chad Djibouti operates
within those parameters. For example,
temperature. When you are telling, okay,
I want to be creative, my temperature this, I want my top be to be this right.
You are defining the school
for your chargeability
applying restrictions. Restrictions becomes your what is
supposing lent that you want the output length,
the token length that you want it?
That is a restriction. But what is that you want?
Then you apply those restrictions that constraints
they create right boundaries
for chargeability
to produce relevant responses. So all these are checklists
that you can keep in mind.

Again, these are pointers. Your checklist revolves around all the components of a good problem that we have already seen.

But as a checklist, you can keep in mind. Okay, these are my pointers which I can follow when I'm designing my project.

So this is how you can write a good prompt.

Imagine having a chat with an API that feels like talking to a real person.

Basics of ChatGPT

That's what Chad Djibouti does. Chad Djibouti stands for Chad

Generators Pre-Trained Transformer

developed by Openai. It is an artificial intelligence based on the technology called Transformer,

which is particularly good at processing and generating human like text.

It's a version of what's known as language model

because it can predict and generate

sequences of words

based on the input it receives.

Now let's dive into some of the buzzwords you have likelihood about generative A.I.

large language model that is Ellen's and Jeopardy.

Now let's

see how to give the prompt to our TED. Typically, I will start telling that I want to build a game app that is rock, paper, scissors.

Demonstration of Prompt on ChatGPT

Let my back and be python and friends

be had HTML, CSS and JavaScript.

I want to execute this on an editor,

namely triplet.

So this is our instruction constraint

because here I'm telling tag utility

that I want to develop a rock, paper

scissors game with Python as my backend

and head stamp and CBS and JavaScript

being my friend. So we are setting up context here.

So this is all about what we want to

develop. Next is I want this for the target

audience of college students

and it should be

for all professions as well. So let it be simple and easy to execute.

So this is where I am setting my audience

pattern and overall context is say, pure.

Then let the creation of this

be very interactive and with simple explanation

and a step by step process

which helps me understand

and create these easily. So this is our overall context

that we are setting here.

So I want it to be very interactive so

that it will tell me what to do step wise.

Then please help me connect

both frontend and backend codes easily

and explain everything from scratch.

So this is again an instruction. So this is my entire prompt, which

I want to give it to my TED typically.

So I will copy paste this prompt

to my TED GPP here.

So I will go to my TTP team and I will give my prompt. If you want to know more about

how to write these prompts in detail,

you can pick up, of course,

from our great learning platform. And it is prompt engineering.

So now let us execute this prompt. When I give this prompt. So this is what typically gives me back.

So so let's create an interactive paper, sees it app using flash at the backend

and hits DMC.

Is his job also different than what it's
giving me step by step guide to follow? So I guess telling me to sign in or log
in to the pledge dotcom.
I will show you a demo of how the prompt
decoding works in a moment.
But first, let's go through the basic
setup for our frontend. We are using HIT DML to structure
our game system,
style it and JavaScript to add interactive
would be hence HTML will help us create the grid
for the rock paper scissors game.
See, this will make it look appealing. And JavaScript will handle the game logic,
such as detecting clicks
and revealing winners. On the backend, we will be using flask
a lightweight python framework.
Flask will solve our frontend files and
handle the game logic on the server side.
This setup allows us
to build a fully web app with Python.
We are going to use
ripplenet and online code editor to run, write and test.
R could replace makes it easy to collaborate and share your projects with others.
Now I will walk you through the steps of building the rock,
paper, scissors game.
We will start by setting up different
then files, then move on to the back end with flask
and finally connect everything together. So here I can go to Google
and I can log in and do my account.
Python App using ChatGPT 4o
I have already logged in so you can create
one account for your own. So this is my duplicate account.

Next, this is telling me

to create a new project. So it is telling me to click on the create

button and just flask as the template.

I can go and click on the creator

applet option here. I should select FLASK as an option.

That is your template. Now I have to give what title to it. So what can we give? We can give it as a rock, paper,

scissors game app. Okay, so now create applet.

Now once you click on create it, open a new page for us.

So now here we will be copy pasting. Good. So let's see,

what is that next step? Our next step is to set up the back end

at this python with flask.

So we have to open main

dot bright and copy this code. That is your python code.

So click on copy code here. Go to our triplet. So this is our main dot

b way. So here I am.

Copy pasting my code,

which I have got it from said to. Be. Next.

After the step it is telling us the step

three that is to set up this printing.

So to set up different then

we have to create template folder first. So let me go back and I am in main by now.

And here there is an option of new. So click on this new folder

and name with us templates.

Okay. Now it is also telling me

create a new folder named template.

Inside this folder we have to create

a file name index dot hitch dammit.

So I copy the same name and I go back

and here it is telling me file. So I will copy paste my html file now.

So I have created an index dotted

HTML file. Now, next, after creating it,

it was telling me to add

DML code, so I will just copy this code and I will paste it here and my HTML file

that this index not legitimate.

Okay. We'll see what we have to do next. Next it is telling me

to create static folder.

So I'll go back on the main dot. Be right click on May not be way again.

Create a new folder by name static. Okay. And Inside that we have to create

two files now.

So one is style dot cases

I will name will file as static dot cases.

So now after naming my one file

name as CCIS, add the file name is script dot JS,

I will create another inside this

by name. Script dot JS. I have created two files inside

static folder now.

So now I to access code. So I have to copy my thesis code

and I will go to my CCIS file

and I will pay speak code here next. Next it is telling me to add

in JavaScript code,

so let me copy

the JavaScript code from here and I will click on the JavaScript file

and let me paste my code here.

Next, we will see what it is telling. So step four is nothing

but running the application on the plate.

So we run into our execution. Let us see the main gist of what

these files are now.

So this is the python file

that are setting the actual context that it is actually playing

rock, paper, scissors and all of that,

taking the input from the user hedge

DML file that. It is used for structuring your code
so your entire web page
will be set up using this HDMI structure. Then as your CSS is very or style
ing your entire web page,
and then is your JavaScript,
which helps you to interactively connect everything on your code
that we are developed.
So now once you understand
it just about it, we will move on to our step four.
That is your running
the application under the blade. So to run we have to just this run button
on top of the zipline.
Then you can access
the game as what it is telling. So now I will go here and I will run.
So when I run this code, you can see that
you have got an output.
So it is dealing rock paper sees. So if I click on rock it'll tell
you does rock.
But computer also tool is rock,
so it does that. So when I click paper you adjust
the paper, but computer rock, so you win.
So this is how you can get the output
for the single player computer being your another player.
So now we will actually define our code. We will refine our code of using time
to be better for two players.
So I am just giving the prompt as drop paper should be for two person game.
So when I give this prompt here again
it will refine the code. I need to give me the code
so some label copy paste
all code of python HD MLC as is Java from here on this same document
that we have created.

So copy the python code. I will go here

and I will paste it in my python file.

So select all select week. So I pasted it here. Now next, let me copy

its DML code so copy could go back to on the plate in your head DML file.

Let me paste the heads DML code next. It is giving me this. It's gone. So copy the code.

Go back to the plate, go to your CSV file and paste your Cisco.

Then go back to your to ATP and copy

your JavaScript code.

Get back to our template. Let me paste my JavaScript.

So after doing everything,

let me run my project now, right? So let me run.

So when you click on your run

button again, you got your new output

for two different layers.

So play add one. Let me choose Rock

and it is telling waiting for the other player player

to let me choose paper so it'll tell

people is the winner right

so it'll be player two is not like this.

You can test for any number of options. You will get the output as who is

the winner in particular option or game.

So this is how you can develop the app

using chat.

You would think

as we reach the end of this review, I hope you now have a clear understanding

of how to develop

a simple game

app using strategy Prompt and Python. The key takeaway is to break down the task

into manageable steps and leverage.

The part of strategy would be to guide you

through each stage. The technology has rapidly evolved

and has made
Using ChatGPT 4o for Statistical Analysis
statistical and data analysis
very convenient. How using the most popular
generating tool said Do it.
So let's quickly apply statistics on one of the healthcare dataset
by using GDP default.
So first we are going to upload
this dataset over here.
Once we'll
upload the dataset of health care, then we're going to give a prompt
what fund we're going to give it here
that you can see that now
based on the shared dataset machine, the dataset and perform data
cleaning is required
using Python and display
the output as well with the coat. So when I will run
this, let's see what output
we are getting over here. So as you can see that I logged over here based on the share of
let's see the dataset.
So first you can see the summary
of the data we are getting over here. And then after we are getting
the potential issues, all the issues,
what we are having with the data data
cleaning steps over here to Djibouti photo is showing
that we have to convert the data,
but we shouldn't discharge the day to day time format correct capitalization
in the name column.
So these are all the basic data
cleaning steps you can see. So once it will do all this, it is showing
the data cleaning is dose

and we are getting over here. We cleaned the dataset,
but we are not getting the
over here. Right? So for that I have all the provide me
the code with the output. So then I'll give this prompt over here.
You can see that
we have the code and based on this code we are getting the output.
So this is the output,
the data description. Also we are getting over here
the sample of negative billing amount.
We are getting over here and this is the
complete clean dataset which is over here.
You can see on the screen
now let's proceed with the next step. So the next step is that perform
the feature in training
if it is required and display
the output as well. So We are not only asking for the output.
With that, we want to add it for all
to show you the code properly.
So let me add
one more point over here that the code
and its output as well. Okay, let's run
this broad and let's see what we are
getting is an output to activity for it's not only giving you the code over here,
it's giving you the explanation as well.
That feature rendering involves
creating new features. And one of the basic feature
engineering is and other times
make one hard and coding is the part
handling missing values a part of it. So you can see that
will feature a complete filter indicating
once we perform based on that,

we are getting the output also and once we'll get the output,
then we're going to apply that
statistical test at large.
You can see the feature engineer dataset
has been displayed for your review. If you need for the modification on this,
please let me know now what want.
We're going
to give the last prompt over here. That is, please provide
statistical analysis and what we want
include in that should include
the collision covariance, the features and the statistical test,
which should have the best chi square
and a not give the code separately and
provide the output as well for each good.
So as you can see, I've given the prompt over here
to provide the statistical analysis that should include correlation
covariance between features
and the statistical test. Like this guys. Got it. And when I will enter this prompt,
you can see the output over here
that is providing the correlation,
the output of correlation matrix. And then after it's
showing the covariance code
how to get the covariance matrix
by using see all the function. So as you can see that we are not
getting output over here, we can ask again
the GDP to display the for
the procedure was in the previous block
based that again
it will read that prompt and based on further step
it will going to display so covariance.
Now we are getting the correlated with

that and output lines will get the output. As you can see, it's a lot.
So after that again it stopped. So when I asked again to proceed. Furthermore now
so based on that it started and it start
showing you the various statistical test so you can see that
we are getting beat this squared test
and know our here first is this
what is the meaning of the is why it is use it is also
giving over here and based on our data
we are getting a code and the code
you can see that we are performing details between building amount
for males and females
and at last you can see the output that what is the p state
value and p value.
So this indicates one that there is no significant difference
in the building alone between males and females
because the p value is greater
than 0.05. After that it showing you
the guys squared test.
Also a chi square test
is used to determine what if there are significant association between two
categorical variables here we can see that
we will test the association
between medical condition and best result. So 24 is not only showing you the code,
not only giving you the explanation
of the guys test, but with that
it's showing you the output as well.
So based on the code
we are getting the output that what is the guide to value
and the p so 0.146 and 0.702 respectively.
What it indicates
what what we are getting as an inference. So from here that

is there is no significant
between the medical condition cancer
and the test result. So because the p value
is greater than over here, 105
Now let's move to B and know what best
why it is. You see it is usually domain.
Is there a significant difference
between the mean of three or more groups? So we're going to compare the billing
amount among different admission types.
So the code for that
you can see over here. Then billing, emergency meeting,
elected billing are using the data.
And from the data
we are using the relevant feature which is called admission type
origin equal to one billing amount
and once minister will going to get the value of adjusted and p value,
not what it indicates
that there is no significant difference. Again in the billing
among different admission type
because the P value is again
greater than 0.05. So you can see what here in the summary.
Also we are getting all the statistical
analysis using credibility for all. It's very easy to use
and easy to understand also,
if you really want to understand. Furthermore, you can give the prompt
instead to due for or
and you can get the relevant outcome. Will be using SDM.
Demonstration of Prompt using ChatGPT 4-o
See this and JavaScript
as our basic languages. The coding background of choice
will be an online editor one compiler,

but judging by default by our site,
creating our dream is as simple as describing
what we envision,
whether it's a portfolio, a blog,
or an interactive application. Jeopardy
for to guide us through the process,
ensuring we achieve our goals
efficiently and creatively. Let's do it first.
Let's open our to activity playground
and that we select for all.
Now let's see what will be our prompts. I want to note a portfolio website.
Let our friends and be active. Look, see if this and JavaScript
I want to execute this on
an online dot, namely one compiler. This goes on instruction constant so that
I want this for the target audience
who have zero coding knowledge are at their learning phase like
college students and only professionals.
So let it be simple and easy
to execute and understand by them. This sets target audience.
Don't let the creation of this
be very responsive, creative, and with simple explanation
and step by step process
which helps them to understand
and create these easily. This sets context to attack DPD.
Please explain everything from that we are instructing tag DPD how it does do.
Give us the steps. Great. Now that we understand a prompt,
let's execute this on to activity four
and see the results. Well, we got the results so fast
and in that structured manner.
For more information
on Tattooed beauty for all, check out our video on chat DPD

for all of us is 4.5.

Great. Let's quickly execute to see our results. How our website looks like

we just copy pasted three different codes

Portfolio website code execution

executing this code is very easy

on the online run compiler platform.

We just have to copy paste

the relevant code in the sections. First goes hits demo, then comes.

The aspect is type. As a third box,

we just copy paste the JavaScript For those JavaScript container,

you can see the watch at extreme right.

You have a preview window. Once you click on the run button, it will execute the complete code

that you have provided

and show your website

how it will look like. This is basic structure

of your portfolio website.

You can enhance the design, color, budget responsiveness and add view of the actions

as per the requirement password.

Thanks for joining us on this journey as we build a portfolio website

with such a beauty for all.

Remember we covered creating

a portfolio website using DPD for all added customized designs

and basic functionality.

Now let us understand what is GitHub,

because I'm going to talk about GitHub globally, so you should understand first

GitHub many you would be familiar with.

Where does GitHub might be using it

also on a day to day basis. However, as developers or software

quarters, they might be using it for sure.

But again, you need to understand exactly what GitHub is to understand

GitHub Copilot.

Before understanding GitHub,
you need to understand what is get because you know, GitHub is
based on the foundation of good concepts.
So it stands for Global Information
Tracker. So get was initially developed
to do version control system
and it's a distributed architecture
version control system, meaning suppose you have developer software or a project,
not that was the initial budget.
Now you did some application,
you know, some functionalities were added or maybe removed and the next version
of the software is needed.
So some
the older version should also be stored, they should be logged upon and then
the newer version should be installed.
So how or where you can do all these good,
provide that platform. Okay, so it's a distributed version
control system where you can keep
your code. That is, you can store your code,
you can manage it, you can modify it, updated
and release the new version of.
Initially, it was created by Linus
Torvalds, the creator of Linux Operating System Kernel as well,
and this was founded in 2005.
Now I talk about some functionalities
of get these functionalities and also carried
forward to the GitHub concept.
That's why it's important
to understand this here as well. So it has a distributed architecture.
It means what that so
there is an architecture there, it's called

centralized architecture where, you know,

suppose you have it in a soft record

and a distorted one place, you are able to access it on day to day

basis by distributed architecture,

meaning it is stored a place,

but every developer's machine or a laptop is able to access it or keep a copy of it

on their own laptop as well.

Okay, so that's your distributed architecture

provided by Get to Give. It helps you run a parallel version.

Now suppose that two three developers

are working on the same of code

or same software project

and they're using it to store

that software on a project. All three developers can run

or do the modification of that code

and the local machine

and keep on doing the operations badly. So the same group will have three

different versions,

three different developers. Computer power really does

maintaining all these versions

so that the three developers can work

independently on their machine for the same piece of code.

Next, it maintains history. So all the meta data related to an all modification adaptation of a code

like who updated the code?

What was the changes done?

When was it done? All these metadata is being

is also being maintained on your platform.

That's what it is. That maintains history. Next is collaboration

and conflict resolution.

So how is this,

you know, doing conflict resolution? So you these of course, as I said
in an example, that three developers
are not accessing and updating it now. So with all three did the modification
in the exact same piece of code,
this code will be stored
back in the original copy. How to resolve
this conflict is also handled by get.
You know there are different logics behind
this. Just used to take care
of all these conflict things next.
But coming to performance
while you're done is fast and secure. I said first because, you know, it
provides an opportunity
to access it on their own machine
and that's how it is faster. And it is faster because it can handle
even large code base.
Even if your code file is very large,
you can efficiently handle it. And it's also secure. It's open
source means it is freely available.
So not just no, GitHub is really access,
it is really accessible, but also you can access the code
or manage your project freely.
There is no such question
then integration. Good.
Can we integrate due to it various ideas.
It can be integrated with various online platforms like GitHub.
So have GitHub comes in picture. So GitHub is an online platform
or I can see a web based platform for it.
So all these functionalities
which are just now listed down for good is applicable to GitHub as well.
And apart from this, GitHub offers

some more functionalities and more features
as it is a web based platform
and it provides
version control for various systems.
Now coming to where does GitHub go by? So you want an idea where does get
where does GitHub talking about GitHub
copilot Not so as GitHub, a web based,
you know version control system,
which was already allowing you
to store your project, your software could basically
and access them also freely.
So associating you know copilot with
GitHub would have been beneficial because a lot of training
data is freely available
and I see so it is open source
the freely available to train the model.
So as I said, GitHub
copilot is a developer tool, so it is an artificial intelligence
based developer tool.
So the programmer is the software coders
or the developers are going to use as GitHub copilot
to write down a code to create a software.
It is based on generative API model,
which I spoke about and it is developed
by GitHub Open and Microsoft.
As I said, initially it was founded
by GitHub, an opening, but as Microsoft is one of the investors that the name
is associated in the founders
talking about
some of the features of GitHub copilot. The first one is they provide shop
an intelligent code suggestion,

meaning when you use GitHub copilot,
the moment you start writing code,
it will help you do you know, So just do the next line of code
or the next few words of that quote.
For example,
suppose your reading of our loop. Okay, so far loop. If you write just the key keyword,
it will automatically
give you the brackets the or know what has to be written inside that record,
those calling records and everything.
Now this example I've taken it from Java. If it is a python call,
it will suggest you accordingly
or it is any programing language
could even suggest you are okay. Next feature. Is it autocomplete the code?
Now why the software
developers are loving this GitHub copilot Because for I think about an on
Java programing language
like I am a dog programmer. So when it comes to job programing,
there are so many you know
syntax is which has to be keep in mind
that is semicolon the call you
you know that
on record so eventually record will be their very own record will be
that should open a semicolon for loop
I should open it after golly records
there is no semicolon required. So these kind of lot of small,
small things are there,
which has to be kept in mind and very sure that other programing languages are also
having these kind of syntax to be learned.
But if you're using GitHub copilot,
you need not learn those, but you need not Google every time
after writing a single piece of code

because GitHub go
by automatically complete your code with the right syntax
of the same programing language
which you are developing your software
or writing a program. It goes faster.
Why does it go faster? Because as I said, the if you're writing
this, the first few words of the program,
our first few key
words of the programing language, it is going to suggest do
what should be the next two lines.
And you just have to,
you know, press tab or enter and it will
write on the code in a faction, I think. So that's the reason it goes faster.
So it understands all the file types. If you have created a file,
you're using it have globally there
and if created a file, the don't by extension, it understands
that it's a python board file.
If you have created of it
dart java extension, it will understand that it is a Java fight and similarly
for other programing languages.
And it will suggest to you the code
or the next few lines of your program on the basis of the programing language of
the file that you have stored that file.
Okay. So that's the new feature that understands
file type, next cloud, good understanding.
So not just, you know, the basic
programing concept and also has concepts of cloud
of suppose your writing or developing
a software that uses a cloud platforms that maybe using any cloud based platforms
it wants to just do according to that

the next lines of code. Okay, It also has database understanding. Maybe any database, MySQL or Oracle,

you will always be a squirrel

or a MongoDB depending upon which database

you are using. It will suggest you are.

I can say it to help

you write the query in that particular database. Programing language, ID integrations.

It can be integrated with a lot of ideas,

integrated development in moments. Lastly, but nonetheless

it gives you the best suggestion

as this, you know, GitHub mobile. It is based on I you know generated model

and this morning this train

on a lot of freely available codes that is open source available code

and hence it provides

you the best suggestion possible for that particular piece of code

which you are trying to write or a program

that you're trying to write that are, you know, disadvantages to it as well,

that are dogs in little bit of time.

But this is all about the features

of GitHub copilot. So what all you can do

with the help of this

you will get a more clear

understanding of this once then see the demo of a hands on session

that how exactly as you know

it is auto completing the code

which you're trying right right or how is it understanding

even where does that in your thoughts.

So let me be very clear here

that it will not read your thoughts. What's going on in your mind, Right? Yeah.

What you are trying to write from the keywords that you have already

written on the get of corporate platform,

it will help you
to write the complete substring. Now let us talk about
what are the languages and ideas
that are supported by GitHub compiler
for getting started with hands on sessions of great copilot.
You need to know what programing languages
you can use on GitHub.
Go find it. So the list is in fact
not limited to this. Python JavaScript, TypeScript,
Java, C-sharp,
C++, Google, Ruby, BHP and many others. And this save the list is not limited
to just these programing languages.
Some might be added more
or it can add in the future initially or the language
on which you know, GitHub
copilot was developed and the first test
was done was JavaScript. So later on you will see that you know,
the JavaScript programing language
you can utilize all the features of GitHub
go by. But I see all the, all the features of
GitHub copilot, meaning different versions
that are different and also parts
are there for GitHub compiler tool that you can fully utilize
when it comes to JavaScript
programing language, but
the other languages are also supported. So this is the which you can go through
and depending upon the programing language
you want to work upon,
you can just get a copilot talking about the idea
supported by a GitHub copilot.
So from the official announcement

on GitHub copilot website recently,
you can see that it mentions that GitHub
copilot integrates directly into your editor, including Neowin
Jack Brains in Visual Studio.
Visual Studio. Good. Right. So these are the full ideas,
integrated development environments
on which you can use GitHub. Copilot I will be using Visual Studio Code
to demonstrate to the code
and also look goodies,
but why I have selected this. So there is no, no such strong reason.
But yeah, initially
when the GitHub copilot was developed it was tested
and developed on Visual Studio Code.
And as you know,
the first always gains of reference. So the reason I'm using Visual Studio
and also as this
you know GitHub copilot was developed
using this tool initially it will be a fully fledged supported
or seamlessly supported
by Visual Studio and that's the reason
I love using visuals real good. Going further for demonstrating
the hand of vision, let's talk
about the advantages and disadvantages
as I have spoken a lot about advantages
they like, you know,
it helps you to develop and write. FOSTER Good,
because you don't have to, you know,
now keep remembering this months, months indexes or as was one
key words of the programing language
get a copilot can suggest you
what should be the next few words what or what

is that one syntax that you are missing?
Why is your code even giving an error that
also can be pointed out by your copilot? So all using all these features
you are able to write the code really
quick, even if suppose in the comment and just as a comment of any programing
language you write, know what
application you're trying to develop, it will give you a basic
good or basic idea.
Suppose you want to write our program
for linear search. You just write in the comment program
of all linear search
and within a fractional segment
that is going to demonstrate to or give you the code
snippet for a linear search.
If you want to develop duo, you want to
create a graphical user interface that suppose to textbooks
as to buttons, to labels, right?
Just just write down this as a comment. It will give you the basic idea. I can say
it will give you the complete good
offered you are you want to develop
that might need some customization depending upon what exactly that you are
you are to look like.
But yeah, it will help you
to write down the basic at least. But there's a disadvantage to assessment.
And they said bus thing is that no GitHub
copilot is not going to read your mind.
It's not inside your brain so definitely
you have to give a proper prompt.
You have to properly write on the comment what you exactly
one good have copilot to help you. It
and as it is trained

on the freely available programs, it might give you incorrect result

sometimes like suppose

I said I want to write a program

for searching or element in an edit.

Okay. And to search an element in an edit. Your copilot has suggested you are a logic

is somewhat like linear search.

Where do you want to do right binaries? So there's an ambiguity here. Correct. Suppose you want to do graded

UI as a give example, right now

and there you want to text boxes and do

what is it right to do the text box. But it didn't grow.

I do believe it thinks the textbook

should be associated with a label. What exactly has to be entered into the

text box is provided as a label, right?

So these kind of incorrect results

might be produce, right? Good. Have copilot.

So it's not exactly incorrect,

but it is incorrect when it comes to the reference of what exactly

you wanted to write code or code for.

So you have to be a little bit

of intelligent programing software. So you need to have a understanding

or a knowledge of the software

programing language

to just check and review that we're never. I know what suggestions are given

by GitHub copilot, Is it right or wrong?

So that's the smaller disadvantage, right? Yeah. As I said,

if you have knowledge of it you might not

be affected with that. Is this disadvantage and convert

this disadvantage to your advantage.

Now let us see some hands session for the different aspects of GitHub

Copilot Starting from the installation

to running program
that you do Java programing language and also executing my asking
queries will see everything.
So let's get started with how to install
GitHub copilot and in the process
Hands-on session on Github Copilot
I'll tell you how to, you know,
integrated with Visual Studio Code.
So just go on Google
and then you can write GitHub.
Copilot Okay. And once you press into
the very first link, if you see a GitHub
copilot your of programmer,
you have to click on this, Okay? And then the link opens up,
this will link it.
Right now you can see all the details
about, you know, GitHub mobile.
It is over here you can see the dart
bias file how the GitHub
copilot is helping a small you know gif. Is it that and as a comment
the developer races under domain
whether the sentiment of text is positive
user web service and automatically generates
this code within a fraction of a second.
Similarly, you can see that
the programing language or Python or, you know, Ruby or programing language,
you are comfortable in working part.
You want to develop a software. The same thing applies coming to the different plans
which are available from GitHub copilot.
So one is that you can use a one month
free trial. So to give it a try,

you can go with the free trial
and that's what I'm going to do it right
now, have it so you can click here
and start a free trial and it will give you a one month free
trial that you can utilize.
Or there are two different plans. One is for individual,
one is for business. So depending upon your account,
your GitHub account,
if it is an individual account
or it's for corporate account, you can, you know, leverage this benefit.
So for individual, is the plan A
or as much as by for business. Similarly,
that means that you can go with buying.
It doesn't mean guys are not going to
you know if you are taking these plans
for the individual and trying to use.
I'll show you also how you know
it is helping in writing a software. Then maybe you can get a clear idea
that I should go over the plan or not.
Similarly, you can see
I'll go through this complete page. It also tells you
how down the line that you know how it is
helping the developers
and in the world that is such as on GitHub mobile it help developers code faster,
focus on solving bigger problems,
stay in the flow longer
and be more for the work. But this is some data and as long as data
which has been built up over it
and similarly down, you can see somewhere
it is also mentioned which all
you know. Yeah, so you can see it right

the program for this. But if you want to draw a scatterplot and
JavaScript, this is how it is right now.
You can see just for the blink of an eye and could
you can go with Python and all the right
drawing as kind of learning python
just within seconds. Similarly,
you can go for, you know, other programs
related to memorization
or you can fetch to it. I mean, writing the program
for fetching the tweets can be written
within fractional tickets. So this is the example
how copilot is helping you.
They go forward now
going to how to install it. So as I said, if you click on GitHub
mobile link, that is a very first link
When you try to go pilot in groups
and it you click on start of your trial.
So if you click on this, it will take you
to the login page of GitHub. So you can see here
that it is already logged
so my account doesn't struggle to reenter
if you are not logged in. So on for the first time using it,
maybe you have logged out,
it will ask you to log in as you can. I go back to that page
and I said, If you click at all
you can just before itself,
you can sign in from it. So as I have already signed in it,
showing my profile, I cannot write.
If you have not, it will click. It will show you the option to sign in
or sign up. Okay, so I'll go back again.
They're clicking on to start of return. And as I have not taken any plan

or are not using the account
which has plan associated. So this is a free version which I'm using
right now to demonstrate you.
And in this, if you can go ahead and see
this option of building out and plan.
So if you go on this Land's End users, so you can see right now
they're saying it's due by September
eight because you simply have taken
the free trial version. If you want to buy it right itself,
you can go here and do the gradation
and you can see what benefits of that
and what is not included. So these are many things are supported
for right now in my GitHub copilot
No goes against your provided
it has never done GitHub. Now how to link it
with Visual Studio code.
So I'm going to use an ID,
but Visual Studio code you can use all the other IDs also,
which I have mentioned to you,
which they talk about pilot support
and the reason using I'm using it have go pilot with Visual Studio
is that because that was the platform
an idea on which to go pilot
concept was initially developed. So for that you should either have your
studio installed in your laptop or system.
If not, then again, you can take help of Google
and you can just directly type.
You're Visual Studio Bird
and the very first link here,
then click on this. It will take you to the page
of Visual Studio.

You can go to this download option
that and it will show you the option if you are using Windows Mac,
depending upon your one, two or whatever,
that whatever is your operating system,
you can download it from it. If you click on Windows, it
it will download you see. Thanks for downloading
Visual Studio code for Windows. Oh, actually I have already downloaded
installed the ID
so I need not install it again. But guys, it's very simple
so you can run to download.
So still downloading. So when to download
you had to just click on it. Click on. I accept the dominance of Aziz
and just clicking on next next next.
It is going to install it late
so it will take 2 to 3 minutes for you to install it
once it's good to download it. Now.
Why does how to connect Visual Studio code
read on GitHub go by? So let me show you with that.
As I said, I have already installed
Visual Studio Code, so here it is, the dominant footage.
So once you install it and open to
your record, you will see that this something, the state of the screen
will come in front of you.
So I have also created
a project folder and a file that the reason is showing it,
but you have not and you have open
via support for the first time. These options will not be
you have to create a folder.
And if I separately later on. Okay so you can actually go from your gear
new file or even open some folder
and then inside that you can do to file

to these options ahead, which you can do. But before that, how to integrate

now GitHub copilot left be as code.

So for that

you have to go to this extensions option also of the marketplace.

So here you have to search the extensions

in the marketplace so at this right

GitHub copilot okay now again

GitHub copilot has various versions.

Okay so GitHub mobile in GitHub

mobile it GitHub will be led night,

different versions of that

which has features and also some modules for different

activities, different kinds of activities

that you do software development depending upon your use,

you can but what I'm going to just show

you the very first basic version of it

that is GitHub 4.7 you right there. The very first option, which you see

with a tiny I of kind of an icon, you can

click on this. So once you click here, like as I have already installed

for a faster demonstration here,

but if you have not in place of this

uninstall, it will skew to install. You just have to click that

install option.

These reload require

disable option would be that there would be only one option given. Do you install it for the first time?

If you have not installed, click on that and will install it

over the fraction of second. It won't take you to any other page

or anywhere it would simply install and.

Then it is going to pop up

asking to authorize be as good by GitHub

to use GitHub equivalent
authorization from as required. So if it doesn't automatically pops up,
then in the notification center
you can click it. And definitely here
there will be a notification saying kindly authorize GitHub
mobile to be used in the escort.
Okay,
so to just click on that notification, it will take
you back to your GitHub account
and it will ask you here to muted bubble,
but says you are seeing that our editor being scored
is trying to access
GitHub account
so you have to allow it and authorize it. Click on authorized button and that's it.
It will link that and once the link is completed to it,
you can see this more like a it.
So this is the icon for copilot. Okay get up open. So when this is active it
if it is not installed
properly it's not active
there will be this icon with up and I'll get makeover it like a slash
on top of this bill
like it means it's inactive right now if you want inactivate or deactivate
you don't want to use it right
click on this and then you can click
on disable globally. So disable globally is for disabling it
for all the programing languages
which you are using on this escort. If I'm using some programing language
subrosa,
then it will ask you that one to disable
for only Java. Okay, so depending
if you want to disable it for some time

or you are finding it irritating
and disable it as well, but would be fine.
But this is so it is closed. And if you are using if you have to run dark audio.
So we as good supports
our programs. Right. So You just have to install the JDK
in all machine
but make sure that duty gives and start. If not, you can again install the JDK
or the extension for Java from here
and install it in the same way. How I explain you forget to provide it so
you are able to run dialog programs. Also
now having to validate have. Okay, so once you have installed
Java you have started applying of GitHub
copilot. You can go ahead with creating a folder
or you can open a folder
if an already created,
but it's there in your system. And then in this you can a new file
and again give it a name.
Suppose full list program dot dot.
Okay so just click into you then
select the folder that's already selected.
We we exclude for little bit
and our project for that already created to conduct the in the file.
So it created
you know the file was programed or made so you can see here so it's a job like
so that was all about your install link
GitHub copilot installing the code, integrating GitHub
copilot with the escort.
No, we have installed GitHub. Copilot we have installed is code and
GitHub copilot with Visual Studio code.
Now time to start writing program the programing language, which I'll be
using the Java programing language.

Let me demonstrate you how to write

on a program quickly using GitHub Copilot.

Okay, so we have created a file

your first program or DOA,

you want to create a new file,

you can go here, click on new file, and then give the extension which you want

to program using Java programing.

So I've given you audio extension. It will you to select the folder, select

the particular follow where you want to create this file

and then just hit enter.

It will create different. Okay. So first program

for Java is a file which I have created. Now you can see your my GitHub

copilot is active, right?

So it's asking to deactivate

it means it is activated to use this. And suppose

going to write a hello world program right

to check

how GitHub copilot is helping you. So you have just written a class

here, right?

And it will automatically give me the suggestion

that I want to write a helloworld program

because for the first time

without any input, whenever you start, the very first thing you learn

is writing a helloworld program, right?

How to print helloworld keyword. So this is what it is going to give me. So I have in class and then I'll just

half of this post it is automatically you can see

showing me the suggestion for the code

and it's completing the line

that after first I should add program as this is the class name and then all this index object

starting with main string

and then the standard output
friendly. Hello. Everything is being demonstrated here
so you can see who it is.
Giving me this edition
for the entire Helloworld program. So I just have to click or
I have to just press
tab and it will come
the end of the suggestion. Okay. And this program for printing
helloworld is complete,
so you can see how quick it was right Now, if I want to bring something else
here, I can print something else.
Maybe. Good morning. Okay.
And if I execute this
by clicking on this icon so you can see that the program ran
successfully, executed successfully.
And good morning here. It is the output right now. I don't want to write this helloworld
program going to the next level.
Suppose I want to write a program for
maybe printing a pattern. Okay,
so there are different pattern programs
which are frequently in interviews
as well, right? So suppose I want to write a program.
I'll just come out. You can read without comment. Also doesn't make a difference
does that you have to delete
the line after the code is written,
otherwise it will give you an error. So I'm going to see an example.
So in the community that's going to write
program to search an element.
Okay. And then edit hit enter. So you can see it's showing me
what should be my next line of code
to press tab and then enter. Now the next line suggestion will also

be given begins in orders and explain.

It has given me the entire code

which should be there for a program to so channel element infinity.

You can see it's asking enter

the number of elements in there. If also it's creating the element as well. Then it's asking to enter

the elements of the area

and then you're going to the elements to

research which element you want to search. You have to enter

and then zones and one zone

it is going to print of addition, else

it will print the message element notes. All right. So let me execute and see.

So this program is for a linear

search, right? If we execute it into number of elements,

and that is, I suppose five elements

are that number.

Now the element will be searched. If I write suppose 90

you can see the element found it position

five because these are the numbers

which I have entered. 2050 was also an 8990.

So when I do the last one position so one, two, three 590

with the position of value.

That's correct. Right now what logic it is

using is a different thing as I have not mentioned any logic to it

anyway, I just want to source

an element intimidate.

So it's using just a linear search. But if I want a specific, you know,

logic to be used

like I won't go binary, search it, I can delete this

and then I can rewrite the comment

and it will again give me this editions. And you just to hit them and enter.

And it is willing
to accept that suggestion. Right? So you can see it now it is using binary
combining searches.
I do find the middle element on the middle. Mm. Plus plus last by two
and then we start searching from first
till the last element. But always comparing with the,
the middle element. Right.
If it is smaller then we go with the all elements smaller to the middle element,
if it is larger than the middle.
And when we go to the other half of their well this is the binary search
and the correct output
is given or equal to
the addition is given by GitHub copilot. So that cricket was right.
Now I'll show you a program here. Do you know create
a duo with two text boxes
or we can write create a log in
from do I get two text boxes and a button
and I have to just delete all these Z. It is also giving me a suggestion
for the comment, like,
do you want to add some functionality
to what you want to write so you can see
how intelligent applications that
and the user clicks of the program
to check the username and password. You want this also to be added later.
If not, you can just leave it as well. So now is asking that when the program
should be logged,
it says log in successfully. So if you want something,
if the username is this, if I don't want
your username to be,
I want I want username to be admin one. Okay.
And I don't want password to your duty

for five for it. I wanted be one, two, 345. I continue to work it
and then we're going to just press enter.
Okay. Otherwise the program should not. It will start writing the program.
I didn't want that particular comment to be accepted,
so I just press enter. So it is not showing me
that I should import
all these necessary packages
needed to create a uniform. So once done is you can see that
if I pressed enter it again.
Showed me the entire code would be done
by press enter. I wanted to
have just acceptance edition, but you can
make sure that it's correct or not. You can see it is extending the frame and implementing
the interface that the action listener
is creating for text field to label
one password field and two buttons.
Okay. When logging in
when exhibited initializing all these in the constructor for the class.
So all this initialization is done,
its creator defined and now all these text boxes and
buttons and password
field has to be added to the panel. So it's adding to the final loop on alert
and then have to implement
action listener on these two buttons
because it's going to perform some action. We could just check that the code is actually correct
and how it is, what action is performed.
That is button
which log in button is clicked and the username is this password,
is this or just fetching the text from it.
And if it is equal to our admin one. So you see it in the Commodores chain,
but I don't know. 991911.

So even if you make certain changes
according to what you want to write a program, it will be ignored
as an input and password.
I wanted one, two, three, four, five
and then it would say that log in successful as log and say,
let's run this code and see.
So yes, you can see as there is no land
breadth and where to define for these.
Why doesn't I work with or join me
like this? But you can add some piece of code
if you want to it for that.
Also the suggestion will be provided
so name. If I write admin only and I write
password as one, two, three, four,
five and I log it, it is going to say log in fail because the username showed me
and with one.
Now if I log in it says log in successful and see how easily
you have created a log in page.
Do I write within a fraction of seconds. So this was all the help
you can see or programing example
which can be provided
or helped with by GitHub component. Okay, now I'll show you as well,
not just, you know, these basic programs
was also the concept of the code
programing language of Java. That is your multithreading concept.
Suppose I want to create a class
which a threat class basically.
So class post program which extends
thread. You can see it is giving me this edition
and if I press tab I've accepted it.

So you see the basic layout is ready. There's a there should be a run. Whether there should be the start should

be called on the object of that class.

So basically out of multithreading

is really now if I want to do something,

start use it.

I can write my own functionality

inside the run method, right? Depending upon what I want and I can,

you know, call this one method from it. If you want to add one more run method

or you can override run method

and just learn or see the concepts of

overwriting with Run, you can do that as well, but you can zero it as written.

One Modern method. So not just the you know

where to go, but it is helpful

for those who are writing a program. We have learned programing,

but also for them are learning it.

Like you can see here, the very basic

concept of threading Multithreading can be learned

in Java using GitHub copilot.

So that is all about our how is it helping to for the developers

to develop the program in Java?

Now let us see how we can use GitHub. Copilot Let me ask you it

in Visual Studio Code. So just know besides our program right now

moving ahead to my school.

So what we should do

here it is that you have to first install the plug in related to my school

and it would be a code

and that you can do in a very few

simple steps that you can just go to the marketplace

install an extension position.

And here in the search

you can write my school.

Okay. And are the second one this

you know if you see this to this database I can tell you to click on this

and install it as I have already installed

again and showing the uninstall option. Otherwise it will show you install option

here. Just click on it.

It will install you without asking you

any accept, reject or next option and simply install it.

So my in school once is installed. So these icons

if you see are on the left side database.

No I squirrel you know these will not be

until you have installed my script

so once you have installed

these icons will be visibility. So if you click on these icons, which is visible errors

if you click on database right.

No, I already connected it

to the database. So it's not asking me

any option to connect.

But for the first time when you are,

you know, integrating my school with the code is going to above or below

will be very autofill the connection name.

Okay. And now what does that connection

name and password which I'll tell you. So let me open my school for that.

So you can see you have all been my actual workbench

which was already installed on my laptop.

So you make sure that you have my school

workbench already installed, You already logged in

so you can see your two connections

are there locally and sends my s reality

and then it by my name. I have created a connection. You can create a new connection from what?

And then if you click on this you will
be prompted to enter the password.
Right? So you can it open. So do you want no here? When I come to Visual Studio here,
when you will click for the first time
this database,
but it is going to ask you to connect this the export to the database
minus group connection here.
Okay this one. So to connect what you have to do here
to fill few details as that both windows not displaying here,
but because I have already connected,
I just tell you what it is up to. I know you do enter the connection
name there and the password. Apart from that you need not change
any feel that.
Okay need not order
anything and just click on connect. So once you click on Connect,
all the database that you have created
in that particular connection and my skill
workbench will be shown over here. Okay. If you want to create a new
I know database, you can click on this
plus icon and create new database if you want to select
any one of the database.
Suppose
I have selected as employs database. Okay, so these are the tables
which I already have in my database
so I can use any of these. You Can see customers Employee's office. These are the tables.
If I want to see what you know, this customer table is all about
you can click on this dot
and it will show you so you can see it employee's
database. There are two entries here
and then there are these other columns. And similarly you can go to the customers
and you can see that as well.

Christine That time will also,

once it loads, it controls. As you can see, you must know my name,

customer number or contact last name.

And all these details are

that which has already been created. If you want to create new,

you can do that as well. Now let me,

I mean, whatever it to see how you know

your tap go while it is

helping us to run my is great grace so for that you can just

go here and the query also

I will integrated these two files

that's why it's showing you the name. Or you want to create a new file.

You can just press control, shift B the shortcut this will be displayed.

You can click your notebook so you can create or open a new notebook.

So if you open a notebook, it will create a new notebook

if nothing is created, right?

So I have clicked on that. So again, see,

this notebook is created here. You can write your gritties

over here as well.

Or the another way

is that you can click on this query option

and then here

you can click on less Like it later. This New dialog box will come out here.

You can make or write which database

you want to use, for example, use,

or you can just write any name here

as well. The query file in

this is basically okay, so I can write.

Okay, it stopped, but so this is just going to create this file.

And now here you can

first you can write the queries. Now the first query is I have to specify

which database I want to use.
So I'm going to use here Classic
models. Okay. Now you can see that automatically
it is giving me some suggestions.
So this addition here you can
see is from GitHub go by. Okay, But I want to use well and good.
If I don't want to use,
then I can write my own query as well. So suppose I want to select.
I'm writing a simple way
to select everything from A or B, okay, so select start
from search showing me some output. Right.
But if I don't want to display everything from customers
but I wanted from employee's DV,
so if I does this adventure me and option scan from employee's
their last name is something like this.
But I don't want that. Also I just want this much. Now you're going to execute it.
You can see here the output, all the data
which is the ad in the employee's DV is being demonstrated
similarly, you can write a rates on complex gradient, for example,
let's say the query to join to tables.
Okay, So if I write select customer name,
you can see added it's already
customer number customer customarily. So I clicked tab but not select everything
from okay from customers is fine red so I have to join to demonstrate.
It's not giving automatically here
this addition. But if I were to have written as a comment that I want to join these in these tables,
it would have definitely
given me the suggestion. But I want to show you that line by line
how on the basis of the prompt
or whatever it takes to put it in good
get aggravated of to so that if your data are left,

join the due date was great.
So it is in a joint. So I just write in a so you can
see now it is giving me this suggestion
right in a joints on customer
our customers and orders they will automatically that's combining
if I want already will do combine with
I can go to my right that is okay let me just know her different options
and you can execute.
Okay so it says that the customer numbers
the list it so you can see how clear suggestion
is that for even the error
it means customer number should be either
specified from which D well I want it to be displayed on
I industry monitor
for time being. Okay, let's execute the way we can see it.
As demonstrated only the customer only if I would have written customers
Dart Customer name Moto2 definitely
displayed is now
if suppose I warn that query for
calculating sum of all
orders for each customer
or maybe for you can write any query here
depending upon what is your requirement,
you can see it as I've just written,
the simple English statement
and it has demonstrated to me that is our select customer name some quantities ordered
in the price of each as total displaying.
It has to order from customers
in our joint order on customer's door customer number,
it was two orders. Don't order number.
Again,

it's in our joining red order details. The details of order.

The price which we are

calculating is bearing in order detail. Now if you want to execute, you

can execute, but you have to delete line.

Okay, because it was through error. So if I execute this you can see some

by name of total is being displayed

for each customer. What is the going to be ordered. Total price for that is being displayed. So you can use aggregate functions,

you can use inner drains, and without

spending much of our time,

you are able to run and execute queries. The various complex concepts.

Lakes upgrades

can also be executed with ease. If you want to create complex queries

like creating an upgrade that also

is going to help do it. Okay, I'll give you an example here. Suppose I want to create a subcommittee

so I'm writing select start

from Employee's Red Employee Number.

Employee Number in select.

So now you can see

it is automatically completed. I want to write this upgrade

automatically on District eight.

So this upgrade by default it has written

is that select reports do from employees they are repositories now

they meaning into something like a self

drain

or it's referring to I'm trying to search all those employees details.

Who is reporting to someone

from the employee side only meaning the reports do is not have an under right

and then you can execute it begins.

It is executing the software successfully. So I think say that you need to know

the concept, basic concept

of any programing language, language

you are using and depending upon what you want to write,

depending upon the prompt or the text

which are already written,

your local pilot is going to suggest the next line of code

or the complete program.

We saw the various examples for school

committees, my school

queries, and also as a Java program.

Now let's talk about cloud. What is cloud into a conversational model? Conversational model meaning you can

interact, you communicate with this tool,

Introduction of Claude

meaning you will be sending certain prompt

or certain statements. It will respond to you in the on the basis

of what statement

you have given it right. So that is the conversational model. Einstein said.

Artificial intelligence. The concept of cloud is based on,

the models of air,

that is the neural networks and deep learning models

of artificial intelligence.

Cloud is developed

by the company named Entropic. Okay, The people who have formed

or developed this cloud

are the ex employees of the company

who developed our DPD. So this activity was developed

by the companies

Openai on GitHub

and, you know, Microsoft, right? So the X employees of those companies

are the people behind

development of cloud,

and that's one of the major reason it is said that, you know, cloud can be the biggest one, but it doesn't you pretty,

although there are certain features

which are better than you. I talk about the difference

between the two as well

because chargeability is

the most talked about the I do right when it comes to the content

or when it comes to

iTunes. Right. And Cloud

is one of its biggest competitors. So it's important here to talk about

the difference between the two.

So I'll talk about that in a bit of time. But yeah, that's the reason why it doesn't stand

as the biggest competitor for deputy.

So they said cloud is developed anthropic

right? The people from the company Anthropic,

what are the different capabilities,

what cloud can do and has a natural language understanding

natural language

meaning you can communicate

with cloud in your own language that is English or India or whatever

it is your language.

Again, you can communicate with cloud

in that language initially cloud was only available

and you can use, but now it is available.

In many countries. It's about 19 languages. It you in summarization and search,

meaning

you can perform the task like suppose

you have given a big text to cloud.

Okay. And if want to summarize it like suppose you have 20 pages or 30 pages

research paper,

or you have a document which contains

of like ten pages of information. You don't have time to read
all the ten pages.
You can provide that document to Claude
and ask the cloud to provide you the summarization
or the output in just lines
of out of that big text. So that kind of text
you can do with the help of cloud,
you can search any information, any data on cloud.
Again, another capability is
you can look at creative writing,
meaning if you want to a report, write a final year student. Suppose you're a final student
and you want to develop
or write a complete report. You can make cloud do that for you.
Next disconnect. If you are a developer or a programmer,
you want to write on a code or a program
for any given problem statement, you can make law do that for you.
Although the capabilities of,
you know, downloading that script file is not that
or it will not develop,
but it will not give you the output
then in there in the cloud interface. But I will write a code for you.
You can copy paste that quote from there
and based it on your machine,
depending on which programing language
you're using. And then you can run it
and make the code execute right.
So it will help you in writing down. Of course, if you are a developer,
makes a surprising plan, which is one important criteria
of using law.
It is freely available,
although there is a problem of client. Okay. The problem is because there are certain

limitations of the open source
or the openly available
or freely available cloud within that is that in
are the number of prompts is restricted.
Okay. And also a cloud to API. If you want to use API of cloud,
let's do the second version.
Then you have to take the pro plan
out of date, the business plan of cloud. Okay, so that it's restricted
but majorly all the does
which a cloud can do is freely
or open source available to you.
Okay. So that's
one of the important aspect of cloud. The second aspect is that it
supports only 5000 words as prompt,
right? So you can see the number of words
in a prompt you can provide, you can provide like a67 pages of
prompt, right. Which is restricted in other tools. If I talk about legibility,
then it's restricted right?
So here in cloud you can provide as long as prompt, you can design.
Okay,
so these are the capabilities of cloud. So I hope got a basic idea of what
cloud is.
It's a conversational model
like DPD itself. You can interact with cloud, sending it
certain prompt, sending in certain
information to certain statements
and then getting back that is once again, those are these days,
which I just no mention.
It is developed
by the company name Entropic. Now what is the benefit of using cloud?
So the benefits of using cloud is not very much different

than benefits of using any tool.

Audio Assistant. It's somewhat the same. The first one is time saving. Definitely Cloud is going to save you time

with respect

to writing any kind of a content

or extracting any kind of information from a document,

from a book or from internet.

Okay. Now, one of the aspect of cloud

is that it can solve the Internet for you,

meaning if you want certain information

that, okay, what happened in the month of October in 2021,

it will help provide you that.

And the best benefit

or the biggest aspect of cloud is that it has access to the data till 2023.

Even the reason data like to want to know

what happened in that year, 2023, in the month of July or in the month

of August or in the month of January.

It will fetch that data too. So the data is up to date. Do it till 2023.

Next is improved efficiency. So definitely the work that you are doing

that is more efficient

if you are doing it using cloud. The example

which I gave was like writing an email.

I'll also show you how efficient

and nice email right? You might take a long time

even to think in that way and.

Right. Sorry,

that's such a nice email. Okay. That's how it improves

the efficiency enhanced user experience.

So definitely you don't have to always

interact with this tool or with cloud

by physically interacting with the system. Definitely typing

or you know, using keyboard
to provide prompt is one aspect,
but there are other aspects also, like you can use voice control
or you can just speak and interact
with client, right? So that's how it improves or enhances
the user experience.
Next is data driven insights. So cloud will also help you in extracting
information that is based on data.
Like I said,
what happened in the month of July 2020? Do you want to get that information?
It will provide you
that as what will see these examples, how
these benefits are leveraged using cloud
when talk about the hands on aspect of cloud know
to understand the technology behind
cloud as cloud is a tool assistant. So the technology behind cloud is
same as the technology
which is behind any A.I. tool that is machine learning.
Natural language. So machine learning, as I said,
is a concept where you make
the machine
learn on the basis of the available data. So that is a training data on the basis
of which a machine
is trained to behave in a certain way. And once it is trained,
once it is efficient enough,
it is allowed to be used on the fresh data natural language processing.
As I said, the cloud is assistant tool and
hope to interact in your human language
and hence the natural language
processing is required. So natural language
processing is a very broad term.

So behind are
there can be the concepts of know genetic understanding, contextual understanding,
semantic analysis of the text.
So these are all others. One is one of domains
that together comprises is behind this
natural language processing. Okay. So let's also talk about
how does cloud differ from other A.I.
tools like you? Pretty okay. As you said, it is important to talk here
because our biggest competitor is cloud the difference key differences
I talk about.
So chargeability is good when it comes to, you know, you can upload the video audio
files as well, even work on the audio
videos files as well in DA Deputy in one of its feature available ones.
Data analysis by the Iot one down. This is available in DA deputy
only you take a plus plan.
Okay it's available in version of da
do beautiful employed mostly everything is available
free.
Whatever it is, it's capabilities
except that the API concept you can use API of cloud
only in its provision
and that aspect IGP has constraint
on length of the project. Okay.
Cloud doesn't have if you paste
a very long prompt and the cloud you know,
chatbot, then it is going to summarize
into a text document
and give it to client
to write into our deputy. You will get an error message
saying that this long message or prompt
is not supported,

although that is very long. Normally or generally
when we interact with da DPD
that long prompt are not required
by there is the limitation.
The limitation da deputies of 12,000 employed
it is 75,000 so that the difference
now the difference is that
the cloud is able to solve the internet
for you and get the data and extract the data from internet,
but it is not able to solve the data
and the data is available
only till your 2021 in DA deputy and it is trained the model is trained
only to the year 2021.
Right? Even the recent data from
the RE can be extracted. So these are also major differences
between body and cloud
when it comes to programing
or as a developer, if I talk about the difference
between the two, then there is not
major difference. The only difference is that entire of plus plan
you will be to create a complete software.
Then in there it provides you the output. It will help
you also to refine your code more better
in a way. In Claude, you will not be able
to do those tasks, but it will also right?
You acquired a program
or come develop a complete software for you and provide you the code then
and in any programing language.
So that's all about understanding
the technology behind cloud and also understanding the difference between
the two biggest competitors of a tool

that is strategy and cloud. Now let's talk about how can we use brand
engineering concept in cloud branding?
And it is a very wide aspect
or a very domain. There are a lot of things
to be understood that,
but here I only talk about those aspects
which will make you write effective when you are interacting with cloud.
So what is? Brown Basically, in a very simple terms,
if I say what is brown input that you provide to any do
so the brown does the input that guides
Prompt Engineering and Install Claude
air responses, definitely you will provide those
from those inputs to a tool
and it will provide you the response
on the basis of what kind of input you have given or what kind of a prompt
you have provided.
Right. So it is only is not able to understand
what is that in your head, right.
What you are thinking. It will only understand what you have put it in the form of words
or in the form of prompt.
So it is very important
to write a very good prompt if you want to effectively with cloud.
Okay. Now what are the basics
of creating a prompt? The first one will be clear
and specific language.
You have to be very clear
in what you are extracting, what you want,
and the language should be very specific.
No, there should not be any ambiguous
terms. Like, you know,
I want the information of some product

on smartphones,
not like it has to be very specific. You have to define the context
and then the expectations,
meaning a very clear input
and output should be given. Content should be defined like suppose
if you want to write an email, many
millions to be written for an organization, for an NGO
or for an AU,
for applying for a job as a free student
or to a teacher. So what is a context?
You have to specify that clearly
and what is what expectations. Like if you want the email to Britain,
do you want the email be a formal letter
or informal email, or you want it to be
a very linear one specifying every single detail, or you want to be
a very crisp and Gonzalez email?
So what is? Your expectations has to be clearly
specified in the form of prone to clutter.
Ambiguous prompts
can lead to unexpected result. So as I said,
there should not be ambiguous terms like,
you know, I want some information about some smartphones
under the range of 30,000.
Not like this. Some some and some right. There should be specific like this
this companies smartphones. I want information
with respect to its storage space
the size of the camera
features of this way. Now if I give you a tip for clarity
in your prompt,
the first one is specific

input and output. You should specify clearly what is input. What is output you are expecting there

to use explicit language example.

So what do you have? Ask for a summary of a book to eyes that you have to specify

which book you are referring to.

Who is the author of that book? Okay. You want summary in how many words,

how many points?

How many pages? Mainly one. To summarize,

the books are within two pages. What do want to summarize it

in just one line

or you want to summarize it in

just ten points that should be specified.

Also, you need that summary of a book. Are you a student? Are you an art teacher

or are you a developer? Why?

You want to summarize the words

of purpose behind it? That context should be said properly,

and you should use the adjectives properly

which you want in the form of your output. Like I want somebody be concise. I want somebody to be,

you know, effective.

I want it to include certain key words,

like I want certain

is you optimized viewers to be used,

and that should be reflected

in their time. These or all these point

or do what you want should be effectively mentioned

in the form of prompts.

So I hope it is clear that you can apply this small concept of brand engineering,

which I talked about.

Just know

when you are interacting with God. Okay, so once again,

what are the tips for effective prompt?

The first witness Be concise. Now they say
you have to mention everything. What is input? What is a context where it is output
certain objectives If you are to add
if you want in your output
to have that features for. But at the same time
you have to be concise
and you don't have to write like
a whole story behind to be very concise. With What do you expect from your
to use correct grammar and syntax?
No, although even if small spelling
mistakes are there, the client is able to figure out how to understand
what exactly trying to talk about.
But I feel using any syntax
and trying to get information from cloud, then make sure those indexes correct
when it comes to programing.
Okay, avoid leading or biased language. So the bias language should not be used.
For example,
what are the key features of this product? Right? You're talking about key features. Okay.
This version like this
is not referring to anything, right? So what product I do expect talking about
should be specified
in what context you are looking for. I looking to buy that product
or you are looking for
in all beating someone about that product. But if you are trying to give further,
give a review of that product.
So that context should be said. Those are the concepts of brand
engineering that you have to keep in mind
while interacting with cloud. Now let's started with the hands
on part of blind,
meaning I'll be demonstrating you
how you can make use of cloud for different set of task,

but before that you to install cloud.

Or maybe you have to understand

how to make use of cloud. So let me tell you here that you don't need to install anything

on your laptop computer.

It's a web based interface. You just have to log in to the account

and make use of cloud.

So you just have to

go on Google and type cloud. You can

see the very first link from Entropic.

So this is a symbol of entropic

the company behind cloud from the cloud. Okay, So you can click on this link.

So You can see I took you to a page where it says daughter

Claudia to enter the email address.

I'm going to do with email

or you can continue with Google if you have already logged

in, you can see your Lord for business.

And then constitutional A.I. is another thing. So cloud for business is mainly for

using API cloud to APIs

if you want to make use of it. So it is available in the pro

plan of cloud to late

free version is available from here. So you can continue with email here

or you can continue with Google.

So I already logged in

so I will not be using this. Continue with Google

and I'll be using my account here and it has taken me to my dashboard.

So this is the welcome screen

that you will get once you log into it.

But if you are a new user, I'll show you. In that case, also what should be done. So I load from here,

you get into the email address here. Suppose I if I enter email address.

Okay, so you continue with email. Are you going to log in

to your email I.D.?
So you either log in. Gold was sent to my email,
which was logged into my phone. So I'll into that.
Okay. So after that
you will be getting this page where you have to enter the full name
or your name you can write.
Okay, so my name I'm writing Trigger here to agree to this and you have to be at least
18 years old to make use of this.
Then click on Continue. You have to also enter
your phone number and get verified
that. Okay,
so the other verification code was sent to my mobile.
You have to enter that, verify the code and then you're
all set to get started next.
So these are a few information
that will be given to you. It may display incorrect
or harmful information. It is not intended
to give professional advice,
including legal,
financial and medical advice. Do not rely on cloud
without your own independent research.
So this is what I was talking about,
ethical consideration. So this is one of the limitation here. Do not make use of cloud,
but any kind of medical advice
or any kind of legal or financial advice, because there are chances
that the cloud can give you incorrect
result as it is trained
on, freely available data on Internet. Okay.
Okay. So these are again, few things that you
have to drive related to privacy policy. And always you have to accept.
Also, there are certain notifications

from Claudia that we major in the uses limit functionality or policies as
we learn more about how people use cloud.
So there is already uses limit to it, but already
you can only use this many front there. I like dislike buttons for the responses
which are being sent or generated
with using cloud. You can like or dislike them
or you can send them back. So this is that.
Then click on finish and here you are. Okay, so you can start a new chart.
Here are all the charts
which you will be doing with. The client will appear over here.
As of now, this is the new account
we just now have created. So there is no history
really to promote users of cloud
that the reason nothing is given under try these I given this example
like summarize this of document.
Help me practice my Spanish vocabulary,
explain how this python game works. We add these kind of single line drums
also can be given to cloud
and you can get that response. You can see it. You want to go to Cloud Pro feature,
you can unlock with control.
You click on that, It will take you
to the billing aspects of it and it will tell you what all benefits
it has so you can
Level of the cloud uses priority access
during, high traffic period. So it's between different versions
of cloud only access to a new feature.
So cloud earlier version was 1.301.2
and then the two wasn't.
And then there were certain worldwide
and before that as well. So if you want to switch
between the versions and you can do

or when you have a plan of cloud the world, subscribe
now it will take you to the
billing aspects of it so you can see it. Cloud Pro is $20 per month plus taxes.
So if you are or you feel
that you have a requirement of it, which majorly the difference is that
until unless you are very much
making use of cloud for really in an order
you need very high interaction with cloud.
Jordan will prompt have to be given
when I talk about number of front has to be given then you can go to plan
else if you are having certain
limited number of branch
but are to be given to cloud, then it's okay to be use a three version of clarity
and second aspect to it.
If you want to use the API cloud two API,
then you have to take this plan
or else you can go
with the free version of cloud it because I'm not going to subscribe it
over to cancel it.
Go to our page. Okay. So you can see here
there's the attachment button. So it says that you can add the content,
you can load the files,
So maximum five files. So multiple selection facility is also there in cloud,
which is not there to remember in our DPD.
Multiple files cannot be selected
one way when you have to select the files. So five maximum files
you can select an embed
should be the size of the file
it accepts PDF text CC. So you can see all these,
you know, document kind of
the text on the good documents are only can only be uploaded,

not the audio or video files.
If you click here, start a new chart
and get started with cloud. So cloud by entropic,
this is how you will install cloud
and get started before going ahead
and interacting with cloud. Let me tell you what our task
we are going to using cloud,
Hands-on Claude
so we'll be talking about
how to generate a good structure. Then summarize a document,
create a report format, write an email,
do a market research and data everywhere,
and write some software goods and programing languages
like Python in Java.
So let's go to cloud
interface and get started. So the first problem, which I'm going to give to a client
is that I want to generate a code
structure for a code
or a subject database management system. Okay.
And suppose I am a teacher or faculty
who is going to teach this subject database management system
to the students of takes?
I can do it and I want to cover
this subject in 50 lectures. So all these is my requirement, right?
I have to put all these requirements in the form as an input
in the form of prompt for class.
So let's write the prompt, generate elaborate code structure
for the subject database management system
for a student of the so so you can do it.
Okay. This was my requirement. What else is the requirement? Okay.
The course structure should cover

all the aspects of the subject
so that a learner can work efficiently
in the real world. One. Again, this is the main motive,
meaning why I'm teaching the subject
right as an instructor. What else? So also I want to complete this course
in 50 lectures.
So the course should be completed in 50 lectures
so you can see the problem to it. It has to be very elaborate
at the same time, concise
and very effective. Okay, so let's give this prompt and see
so you can see that the speed by which the client is in writing
responds is not very slow.
You don't have wait much like in for example,
if you have use Googlebot that does speed
or the speed of response generation
is very slow to wait for some time. This is equivalent to charge abilities
page version.
Although the free version of to
our deputy is more faster than this ready. So you can see how beautifully
it has generated
a good structure for the subject database
management system. So here you see here is a proposal
structure for this introduction
to lectures
and also divided into number of lectures each topic should take to cover data
modeling, school
based design,
and even the advanced concepts are covered because I have eyes
that the course should be enough
for a person
to work in real world on database. So yeah, it has taken all the advanced

concepts into consideration.
Also the programing aspect of database,
like how you can connect to these and all. Is it also important
when you're working in real world, right?
So all that is there query
processing and optimization. So yeah, Lord has clearly understood the context
which I have set up for it
and has given me the response accordingly. So this is how you can generate
the responses.
The first example we saw. So the next one which you are going to see that
we are going to summarize the document.
So for that I'll be uploading a recent article just downloaded from Internet,
which is really available.
The research document is on digital image
processing, so I'll just upload that at
the size constraint. Is there that an image should be right
and it's not embodied enemy. So yeah, this is the paper.
Okay, I'm not giving any from that,
but let's see what led us to this. Okay.
So you can see ahead is an edited and formatted
version of the research paper. This is a research or digital image processing technology
and it's applications by this person.
Okay.
Okay. It is providing me
the complete detail of the paper right.
Okay. Let it do test it. Then I'm writing the next prompt
that I want to summarize this document.
So okay,
so I can use the polite words, right? It is more interactive in that way.
If you are in polite words
like please kindly and alert. So I'll use please provide summary

for the document uploaded above

in just ten points and just ten points.

Okay, so this is what it has done. It does read out the complete paper, but not completely in a concise way

or in a summary.

This is what has read it, picking

two aspects under each subheadings

I feel again. So what I wanted

is to somebody for the document. So document uploaded above and just input.

Okay, let's

see what response is being generated.

So he has a ten point summary

of the research paper which is uploaded. So this can be any document, as I have taken an example,

randomly picked any research paper.

So you can see the end point. It does give it the just covering

everything which is there in the paper.

Yeah. If you want to know summaries

the document in just a line.

In just one line

so you can see in a single line. Also it can summarize that huge 7 to 8 pages of all of video for a document.

So this is what is it

somebody or the output? So this is how you can summarize

any document.

Next, let's see,

I am going to be generated an art. So for that I have a meeting. Transcript. I'll upload that meeting Transcript. Here.

Okay. And then you can see how that can

you please

generate and further attached Meeting Transcript.

Writing in room is a task whenever there's a meeting of the people

who are working might relate to this.

So if you have a meeting transcript

I don't see written document or a meeting transcript as a whole.

You can give that to a client

and ask it to create an email so you can see the date, time, location.

A proper format has been

for creating an email with what attorneys and then words agenda

and then what are the key

discussion points

and what should be the action like? This is a proper format of them should be

right and that's ordered as followed.

So these are a few of the aspects

how we can interact with cloud. We'll explore some more

benefits of cloud and some more prompts. How in response to that. Okay, now let's see some movies.

How you going to interact with cloud? So we have seen three or four aspects

right now.

We'll talk about how we can make use

of cloud for writing an email for us.

Okay, so the context is that I have a

email for applying for a sick leave, okay.

For two days and you have to alter

the law that you're working as maybe a developer in X, Z company,

or maybe in any I.T company.

And I want to apply for a sick leave

for two days because I'm done with it.

So let's give this prompt and see how. Good email is written by you.

All the input is given. Let's see how it generates any.

See how you can see how effective and an efficient email for sick

leave was written by Claude.

Read this effective email. You might need at least 5 minutes,

but this did it in less than a minute.
So yeah, the subject the manager name
and that I'm writing to inform you that I'm looking at this from day
to day to the state.
I've been suffering from and Gulf
condition has worsened and also how see how nicely it has framed
both situation and situation.
I plan to use this time to rest properly
so that they can recover and I'll respond back. So yeah, even though nice finishing lines
for an email is given,
so please let me know
if you need any additional information about leave application. Apologize
for the short notice and inconvenience.
Okay. So regards immunity. So this is a very effective and efficient
email, right? Written by Clark.
So and they have already mentioned that
if you want to write, if you are working with content writing and also for
creative writing, you can use cloud.
So if you want to a report format. Okay. Well, for example,
can you please provide me a report
format? The report is for
the final here project of an engineering
college. Okay, so you want a report format
for writing your final project
of an engineering college. Okay, so this is the project that you have
written a little from the very small.
You can elaborate it again
according to your requirements, more So let's see in this
and this is able to generate
a format for me. So yes, you can see there
should be a title page with those name

statements, correct. Title of the

This is how we prepare. Correct. So we're a name in already number. The example this I took because

most of you will be able to relate to it,

but most of you would have prepared

the report for your final project or. Those who are still in the video

call might prepare, right?

So there should be certificate

and acknowledgment. Then. Abstract Table of content List of figures

list of tables,

literature review so you can see you in

point those later subsystems and their

system design and implementation

where it all should come and a result analysis, work, references, appendices,

all that is included.

So this is the perfect report

for my generated. Similarly, you can donate a report format

for any kind of report

which you are looking for to prepare. I bet it's a financial report,

medical report or registered document when it comes to software

development order.

I know what would be the format

of a design document if you had only one. So all these report

formats can be generated using slide.

Okay,

so now let's see the last project video and we have to cloud

that is for doing a market research.

Okay, So we'll be doing a market research

research on a product iPhone 15 and we'll also see try to find out

what is the current uses partner.

Okay. So I'll give a prompt. Can you do a market research

on the product iPhone 15?

Okay Also also give me the latest

uses pattern for iPhone 50.

So this is how you can see

the response has been generated by cloud. Okay. So it has given me the market research

like then

the reason data is collected

and given the response. Okay,

so key markets like this in the worldwide,

what is the market aspects,

as I have not mentioned, but I want the market research only

in specified country, India or anywhere

as it is, and it has given me worldwide

and also uses partner have eyes. Right. So you can see the

whether in map apps uses has spiked

so all these information gaming adoption

rising as the A16 bionic enables

immersive graphics and experiences

so all that has been considered in the latest iPhone model

that is iPhone 15.

Also it does summarize that

strong performance aided by aspirational brand image and key feature additions like

satellite connectivity is the advantage.

Also, although pricing and competition

remains a disadvantage. So this is how glargine rates

are responses based on what brand you have

given. Now lastly, let's see how cloud generates

a software code for you. So if I want, but can you

write a Python program for TIC to keep

Program for tic-toe game using Claude

program to you? So it is going to write you

a program, okay.
But it will not provide you
any downloadable link or anything of that sort
which was available in charge.
You would right Virgin. But it will write you a correct code
and if you will give
the same prompt in other tools,
it will also like other items. When I write it is chargeability
it generate almost the same code,
meaning that client also generate
the most itemized code for you. But it might be there
and sometimes it doesn't.
And that gives you further
give it a prompt to optimize the code so you can make the conversation
interactive.
As here. I have not made it interactive
except for the case of the uploaded a PDF document
where we summarize that first
and then point that in one point
and so on. But if you want additions and certain
things to be removed in the response
which has been generated by cloud,
you can do that as well. It is a direct feature. You can interact with client like
how you interact with any other human.
Okay, we can see this is generated program for a detect. Okay. Okay.
Still generated. So here is a two player game. Okay. One is you
when one is just zooming to the computer.
So you can just copy the code from here
and see if it
runs or not because again, copied from it and you pasted in the ID or
like this is a python game, right? So python codes you can pasted in

any python ID and make use of it,
but executed. Similarly,
if you want some program to be in Java programing language, you can give it
a prompt accordingly as well.
So this is how we talked about
how we can interact with cloud by our understanding how we can give Brown
for different aspects like
genetic structure, summarizing a document
for a report migratory or generating an event
and also writing us off before in Python.
Now let's talk about plot to API. So glad the API is available only
for business users and it's not really.
Claude 2 API
We really have to take the pricing plan
or the plan of plot too. And then only you can leverage
the feature of cloud to API,
although I'm going to talk about it
on the steps which is required. What do you have to follow? How you can write a program,
make use of the API now for your own use.
So cloud to API is being offered
to business customers at pricing plan. As I already told you, you have to.
You can use it through API
as well as python typescript clients. Right.
I'll demonstrate you a program here
how you can write it and leverage the
benefit of to API for Python.
But before that, if you are not a business customer of cloud two,
you can also make use of it,
but only selected customers
are given the benefit of it. Okay, so this is the link
which is provided, right?

If you click on this. So once you click on that,
you will be reached on this URL
that is WWL blue dot entropic dot
com only access so cloud API access.
Thank you for your interest
in entropy language model API is currently being offered
to a limited set of users.
We hope to expand access in the future for you to fill out the form here
which is mentioned below.
Okay. By the form,
the email which you are filling here should be your company or organization. Email address.
Okay, fill the details. Submit the form. Okay.
Once you submitted on the email already,
which you have mentioned, they will provide you the link
to access the client API.
Okay. And that will be an access to the console.
Okay. So once you click on that link
which will be sent to your email,
you will be having access
to the console of Okay. And then you can make use
of the benefits of the API.
Although when I applied to it,
this has been like a month, I have not got the response back
from plot.
Maybe the number of users whom
the access has to be provided is already filled or I've no idea
what's the reason behind this.
But yeah, the access is not granted. But yeah, if you are in immediate need
offload to API benefits, then you can take
the pricing plan of business customer
and then you can make use of activity.

Okay. So once this is done, as I said,

you will get an email, you will have to click on that link

and go to the cloud

console. What after that,

what steps have to be followed? Let's understand that. So once you have reached to the Cloud

API console,

it will generate the API key

that you can do it by going to settings. I'm not able to show you the hands on here

by the steps.

I'll make it very clear so that if you are a business customer

you can make use of it. Okay, so when do the settings

and the user settings

you can generate the API key from that. That API key is what you are

going to use it in your program later.

Okay.

So make sure you do generate is API key. Then you can write the program. The program I like to write

will have these five steps.

The first is you have to import

all the required modules. Okay. The modules like so can see here,

this is the code.

Okay, I'll be talking about the steps

and also explaining the code side basic. So they import all the required modules

like entropic your from entropic

you have to import entropic uman

underscore prompt, underscore, prompt.

Again, this is for our understanding human prompt

and this is for generating a response. Okay. So these are the modules

because you have to import

also import OS module. I think boarding all the required module

is to initialize the client

using the API key

because they have to create entropic link. Okay.

That client is what you have to use. Or you can see it's a kind of objectivity

that you'll be using in the entire to generate the response.

Okay, we have to initialize that client

using the API key. Make sure to have generated API because that is already

being used over here later in the program.

So here you can see in this line of

support the step you have to generate it to use the generated API key

and that is the here key here.

This is double code I've written entropic

after school to appear in a school. Great. You don't have to write this here,

but you have to write the key

which will be generated once have been given

the access of cloud to API.

Okay. Funds are generated

by going into the settings, the API key that has to be used or buried

and then you have to create this class.

So this is a variable

going to give any name to this I give any underscore

client equals to entropic.

This is the name of the class. You can see the module name and inside

this API underscore gig was two was module

environ the smart environment or indeed

the spelling is this only environment.

Okay and in this grabbed package of

to write the key. So this is how you will create

a entropy client.

The next step is that you have to now generate notice once by providing a model name,

maximum tokens and prompt.

So you have to also specify where

the maximum token token is also something

which you provide input like in the prompt and what the model name. So that is what you have to provide

information.

The next step. So in the program is this you can see here

so, so even this is a variable or object

and topic dot completions don't create

and in the square so in their own bracket

to provide them or delete

the maximum token then the prompt. So model as close to this is fixed. Okay.

You cannot change that. The model will always be client hyphen

two maximum tokens. You can change it whatever you want.

So 300 is what I'm going to evaluate. Now the prompt, make sure these key are the same model,

maximum underscore tokens and prompt.

These are not to be changed from should be provided it

human prompt and write the prompt which you have to provide as an input

to get the response from the cloud.

Okay. And then write it from here. So this is the actual statement

input from originate.

How do writer email using gmail account? It is a one line single prompt

I have given, although this prompt

is not very good or very effective. But just for an example

I've put it over to okay,

and then you're to print

whatever is the response you will get. So print completion

don't see this object as everything.

That was a last step, so you have to print the response

before that.

Your prompt usually has human

prompt and prompt. We saw in the code, right? No human prompt any problem. Right.

And then lastly, have to print lotus one. So these are the steps

that you have to follow when you are making use of cloud
two API in the okay.
And this is the program of program
is very simple, very effective, just few lines of code and you will be able to use
cloud two API in Python.
So I know demonstrated
you do the exact hands on session there because as I said, the early access request
has not yet fulfilled by entropic to me.
But if it does for you, you can make right
the same program there and check
how you can use the benefits of cloud two.
Now understand how do we integrate Python
Integration of Python and Gemini 1.5 pro
and Google Gemini 1.5 here? There is no manual work in need,
but there are a few steps
we have to follow. Let's explore and understand
how this integration happens.
Let's quickly
switch to Google Studio interface. First step you have to follow is click
on documentation and clicking on that. It will open this Google API
for developers interface.
Here you have Gemini API, you have Gamma, you have Google
Edge and Tools, you have. Choose. Google, Gemini API.
But by default,
this is where it will launch. If you want to get API, you can go here.
Else I'm going back
and just clicking on the quick step.
Once I click here, you have a quickest way
in order to choose between what language
you are working with. Either you want Python by default
or despite our Node.js go

that that is out for Android

and iOS development.

You have Swift and see. You have already made option

in order to.

Work. With Python. With the help of Google, I studio.

So I'm clicking on Run. Once you do this

you will get to see this interface.

What is this all about

you can see a predefined integration that's happening between Google Air Studio

and the Python interface

that this Google CoLab generally

we use back here. You can just run

install Google generative so it will install all libraries

available

in order to work with generative. The complete package is been installed.

Once you run this, once it is complete,

it will show that

the particular

air packages are being loaded. You are collab see,

it gives us a checkmark.

Then you'll have one more option here

in order to get an API key

to setup this particular call

you have to just run this.

Let's see 100%. You will get an edit. What is that error? You should get an API key and.

You should try to put that and remember the API keys are always important

in order to keep it secured.

So should not share your API key. So let's resolve this error. How do we do

here? You have to click on the key option available.

It is our new secret key. If you don't have Gemini API key,

you have to click on create Gemini

API key. It will end up to get API keys interface off Google Air Studio.
You have to click Create API key. It will show safety setting reminder,
read the dumps and click on.
Got it. You have to just press on, create
API key in a new project.
If you are putting up API key
for the new project and you can search for any old project,
but it is having my first product here.
Now let's choose to have a new project. I'm generating that API key
for the new project so it's loading. It will take some time
in order to generate it.
Once it is generated. We can copy this particular API key. Okay,
you can close this. Go back to your our new secret
in the collab of Google Air Studio.
Click on ADD, you have to give a name,
what name you have to give here.
However,
it is mentioned that this in capital itches Google underscore
API underscore. Key so. This should match
the particular format
how they have given in this code. I've given the name
and the value is what you copied from the
get API key interface of Google studio. Then remember
you have to switch on this button
Once it is checked, you have to run this
particular code once again.
Right now
it will establish the connection. See it says right mark here.
That means it has been loaded. The API key is now connected
to your python
Google collab of your studio.

Then our work is done. If sometimes there comes an error. Even after pulling this particular API key, just go back to this and check whether this toggle is on or not. Every time you use you to have this button switched on. Next, if you want to see the model details you can run. Don't do it. Do more than you can initialize this. So I am initializing what Jimmy and I put up. So everything. Is written. Actually, this is just a sample you to generate a text. It says in the content. Write the story about the magic backpack. So already it has been run. So You have a complete story here. If you want to write your own course, you can use that. So there is no use of running, but Still, I'm running this so it will generate the response after text. However, you have given the prompt for here it is the text generator, a small story. What's next? If you want to learn more, you can learn across in the Python tutorial available integrated to this platform itself. Both the integration of Python and Google Gemini 1.5 Pro is very easy and understandable. Now Let's understand building a simple chat.

Chatbot using Gemini 1.5 pro

But before doing this demo, let's have a basic knowledge about what is a chat bot, what are the types chat bots and major applications of chat bots across domains. First, what does it support? Chat bot is all software which is developed in terms of artificial intelligence

where it can mimic
the human conversation or. It can also answer on behalf of a human
in order to address the critics.
Generally, these chat bots are used across
nowadays integrated towards
each and every software you open.
Okay, go to any website
which traveling, booking website or food delivery website
are any shopping websites,
galleries, clothes,
anything For the motto Ask small chat. It will pop up and it will ask us
for if we need any help. How it is working.
It is having a pre defined trained data like if we question on something generic
which can be handled
and addressed by the bot, it will resolve
the particular requirement.
How it is helpful. It is giving us a support in order to minimize the human intervention.
If you take an example of swiggy, if your order is delayed
or some of the other problems
happen first,
the approach they follow is by resolving these issues with the standard equations
which already has a predefined answer
for that in case. Beyond that,
if you want to talk to an agent
or a customer support person,
then you can choose to, talk with agent. They in a first place, they are trying
to reduce the human resource
or intervention of the human
in order to resolve the issues cost
for a customer from the service
or the product. So this is the approach nowadays

we are following

and we are accustomed and used to in order to interact with what's next.

Let's explore

what are the types of chat bots we have. First, as rule based in rule based,

it is having a straightforward

quick rule in order to be followed to get the answers

we want, as simple as that.

Next, have AI powered, which is a little

advanced watching of rule based. It is having a support of machine

learning in the back end

and it can address the wide range queries

across next guns task oriented

Task oriented as following a certain

set task, the bots will do the same task

monotonous again and again.

So the boredom we have as a human beings,

it will not be felt if we use chat bot

for a routine auto normal repeat the task

so task oriented bots will help that way.

Next comes conversational. As I told you before, conversational. Generally we converge with the customer

support for any of the product or service.

So their intervention of the chat bot,

which just conversational will come across.

Hope we have a broader knowledge on

what are the types of chat bots we have. Next, let's talk

about the applications of chat bot.

Yeah, chat bot is there everywhere. E-commerce online

delivery apps means on Myntra.

Ajio you take any of the e-commerce

places, you have intervention of bots.

Next health care,
any application of the hospitals.
Has this support. Finances, you have any bank applications
you want to log into your net banking,
you can get the option
of having a chat bot or if you want to open an account, you
go to the particular website of the bank,
the chat bot will pop up
and ask what you are here for. What? How can I help you?
That is how will try to resolve
the issues. Other address requirement next education.
Here we are using in all the products we are trying to create and understand.
Be aware of what's the technology
that's making our to create
this chat bot next comes entertainment.
Here there's a broad perspective
bead, gaming, bead storytelling,
having certain visualizations, graphics every where chatbots can be used up.
We are clear with the small introduction
of what this chat bot types of chat bot and applications of bot.
Now let's move on to our creation
aspect demo. Before starting our demo,
let's have some groundwork done
first, how do we design this? We'll keep it very simple
and very comfortable.
Number of lines of code
in order to understand because it is very basic chat bot
we are not integrating too many things.
We are using Google Gemini
and we are trying to integrate that API.
We get in order

to make a conversational code. That's it. So designing
conversational chat bot involves
interaction
between users and the computer. One thing, including
it should have a greeting, it should have
a response to the greeting back
and you should also provide a certain exit comment as well,
which is very much important.
You have to ensure that chat bot
will not end randomly. If you press an option called exit,
you can type exit
it will exit out of the code. You should keep it very organized. We should not up an edit or lander
in middle of the particular code running
and it should not go to error forever.
Right. We have to ensure
it will be a smooth landing and smooth take off
so that's how we have to plan across.
Let's understand the code
and then walk along together. So last I will be showing you a code
when it is not conversational,
but it is serving the purpose,
then we will make that particular code,
improvise that
and make it that conversation checkbox. Let's quickly go back to our Google lab
in order to understand more about this.
Now here we are in our empty collab. I have just named this collab
as a chat bot
that I beat right and beat the Python fight here.
Let's Start with that first line of code. What we should do. We have to import certain libraries
we need in order.
Don't start off so I'm using the keyword import and trying to import.

Don't google generate. I

and I cannot be using this particular whole sentence.

Ask Google dot generative AI. Right. So I will be giving our name for it

as Jenny AI.

So short form of it then what

I should import I have to import up

or else what is this? We are trying to set that environment

with the help the OS library.

Okay. Then comes are crucial part. What is that part?

We have to import our API key.

You all know how to get your API key. I'll quickly give you a revision here.

You should go to Google studio. You should click on get API key. It will land to this particular interface.

Here you can create an API key

or use the existing API key, then paste that key once we start coding.

So let's go back to the coding

platform and check what we can write.

We have to first call the environment so OS Doc and we don

and I'm

trying to create a API key variable. So it is API key.

Then what is our API key? We have to select and paste it inside the single code.

I've got my API key

so I'm pasting that right here. So I well I start environment is

API is also been set. Next. What we have to do is we have to create

the model.

Model variable is equal to

we have to call the djinn AI, which we have already imported

and we have to use gem.

I got generically model so Jen, I don't generate

do you have to write the model right here?

So it is clear now next you have to use a

small bracket and you have to
put up a single code inside this code
what you want to write, which Gemini model
are you using so Gemini 1.5
pro and you have to give a complete detail about it
by adding a word called latest.
The latest version, which is present in Google AI Studio, is what it is taking.
So that is why we have to
mention the same. Next, you have to create response,
a variable
which deals with model interaction model generate
content. You have to then read the content
for the question given by the user.
Right? So model and generate content is equal to
who are you In our double code that is to be done.
Then you have to print the response. How do you do it? By using our brain
function print response,
dot text. Right. So this is a simple code you have to write
across one to do this coding.
Let's try to run this core. Now. Once you run this call,
you will get an answer.
Let's see what is the answer
from the Gemini 1.5 Pro editors.
I'm a large language model
trained by Google. Think of me as a computer program
that really has
a good understanding and generating
humanlike text, so on and so forth. It gives the completely this.
What can I help you with
today is what it is questioning. And if you want to do it again,
you want to change still from here,

which is mentioned in this double code,

which is not that chalkboard

kind of feeling right now. What we have to do is we have to attack

what feel,

we have to make it conversational. So how do we make it conversational? We are just extending the same code

and making it conversational.

Let's check

how it is done before extending this code. Let's check out the output given.

While I was talking,

I just changed as how audio

which was before us. Who are you? Right. So it is telling. I am a language model

so I don't have feelings like humans

do, however, and so on and so forth. It means it is working. The conversation is happening,

but it is not happening in a chat bot

that's quickly give up good optimization for this and extend the code

and make it conversational.

He all the conversational code

that's ready. What are the new additions we have made? It was very easy to come,

so let's do that.

Okay. First line is same that we imported

Google Dog DNA to be a genie.

We also imported the oils environment. We set up our API, which we usually did

in the initial code

in its same configuration. So we are trying to configure

this particular API key

in the OS environment as API. Next, what we did is

we selected the modern generative model

that allows US same line here comes how we make this particular code

which we actually executed into our chat

bot more for that. We need a function chat with bot.

Here we are trying to give instructions

that if you are in case types,

exit or quit

the chat bot should mention it as a goodbye

and it should break the particular loop.

Next, what happens? We have to give our header handling matter

that is using client except here

if the user input is something wrong

which chat bot does not understand.

We have to give up common

printing statement that does. Sorry,

there was an error processing request.

And what was the error? You have to mention that right here. Next, we mentioned our main content

and the check loop.

So main function should be that you have

to call this function of chat with bot

and the loop continues until. The user input will come as exit or quit.

If you try to do something in-between,

it will add up the error loop. So this is not very much

helpful and gives a good feeling

about the chat bot which we created. I hope we have got the extension

and what are the extensions we have

done is been told and it is clear

I'm trying to run this code now and let's understand.

See the first line which we gave here

that is chatbot is ready to chat type exit

if you want, end the conversation.

We are trying to give input to

that user as well where they understand they can type exit

in order to end the conversation they want

you is what we are that does user. Hi, I mentioned this I could on

and so what happens?

The chat bot will give an answer.

Hi there. What can I do for you today?

I said how are you so? Once I do this

it will give a good conversational way.

As a large language model, I don't have

feelings and expand so on and so forth. What we did the same thing? It is just telling, What can I do?

How can I help you today? So next, what is dataset?

I am trying to give this prompt to that chat bot. Let's see.

Do we get the definition of data science

or what it will do? Let's explore, let's wait and watch.

Okay. You could see the answer. It does have very descriptive answer

where it says what does data science,

what are the key ingredients of it?

What the recipe offer

and what are the applications. It is not. The words we just use

is not that relevant, but still

it is giving us certain answers, right? So the applications in natural

what is data science?

It is giving a summary as well,

trying to give us that particular summary. So I just scroll here till here. So it is giving us this complete details.

So I see exit now chat about this goodbye because we have already coded

as if we get exit or quit.

The jackpot should say goodbye. So it does walking grand. Fine. Hope this simple chat bot which is built

is very much easy to understand

and you could integrate that

with the help of the API. Give it are generated.

Now let's start with the first topic overview of Python. When you hear the name Python,

you know the various applications of it. First and foremost thing,
it is a high level programing language
which is very unique compared to other
high level programing language, right?
Almost. It will use English like statements
in order to execute the code.
It's very easy to learn as a beginner this particular python language
now vitally use python in generative VI. It's not about generative AI,
it's about Python is already having a well supported set of libraries
which is already in use since it's with respect to domains
like data science, machine learning,
natural language
processing, deep learning, etc.. Now artificial intelligence
and AI is grabbing the libraries
which we have already in Python are that
programing languages are also used.
But I could say Python is a versatile
programing language which makes life easy for the people
working in this technological domain.
After understanding
the overview of Python, let's quickly hop on to the next topic.
Introduction to Generative Applications. But just a core concept we have to learn.
Generative AI Applications
Then they're going to first do algorithms
which enables machines to produce content that is not only new and original,
but also includes a reflective data.
And it will be always screened
according to the requirement by generative ideas with a lot of models.
All of these models include genes that is generative adversity networks
vs variational auto encoders transformer based models
such as tangibility.

Right? What do we do with this? In the device application? It's very important in order train the algorithm or the machine in order to keep it updated. The more you interact with this, the more it gets trained. That's how simple it will and then do helps you to generate your own models. How you want to train that particular model, you can train it accordingly, just like simple. How does the scientist train the robots? Each robot will do its own different work, right? Poppy would have seen the requirements are different. The catering of requirements is different. Hence the models will be trained accordingly with the help of the entity. Yes, it includes a lot of other technologies, deep learning, neural networks, etc. but still the network is also AI based. What does the significance creativity boost? It creates enhances processes by providing very good content ideas, new content ideas. The new way to the problem efficiency. It is giving a helping hand to human beings in order to be more efficient. The more good you use, the more productive automates. The content creation or saving time is very much important. It's an important resource. Now it h to this particular saving results, then personalization. It generates particular personalized content as per your requirement, as per the prompts you give the tag. That's how it works, right? So it will cater various applications. But the theme this is the overall picture of generative AI applications. Now let's about the next concept

development environment set up.
How do we do that? What is it all about? You have to have a platform
in order to work with
you need to have a basement
in order to build a building, right? So let's learn
how do we build this particular base?
So what does this thing consist? It consists of few steps in order
to set up a particular environment.
It is not dealing
with much higher softwares or something from moon and stars.
It's very simple. You have to go to Python official website and download the latest version,
but now it 3.4.
Well, you can download that python
into our local system and you can execute this via
command, prompt.
Right? First step we have to open command prompt which we have in our local system,
and then we'll have to navigate
to the location where the python is installed.
Both everyone are having a clear idea of
how do you work with the Nixon unit,
at least a basic command like CB change directory, make make directory.
So only these two commands are mostly used
in this complete session. I'm not going deep into advanced level
of Unix Linux and if you want to work
with the command prompt,
you have to just use C.D and make command
and you can make directory or change
directory. Simple as if you're using windows.
You can use command prompt or position for Mac voice on Linux,
you can use terminal.
This is the platform

you need in order to work with. After navigating to the location
where the python has been installed,
you can install all the libraries
using be IP connect, right?
So be it. Be installed as a basic command and you can change the libraries
you want accordingly.
First, the talk about name by number is very well known
amongst the domain call data sites
like the first thing is number
three will always cater in order
to help the mathematical calculations.
Also working with high level
data structures and give you the complete access to the functionalities
and automatic and logics.
That is right. Data science is dealt with a lot of data
numbers and other elements. We use number for the same.
Then we talk about logic. Then you hear the word flask. It is a library
which is related to Python,
that is this web based framework. You can create web application
using this particular flask book
that does the major
help of using flask. The next one is stream or stream l.
I think this particular library deals
with visualizing the models created.
Then have thought, thoughts, vision
and thought audio. Basically, this library caters
computer vision models.
You can work the model creation,
you can view the model and also can add to the multimedia
the model created, right?
You have this bunch library

in order to cater computer vision products models,
except you have transformers.
Next, Transformers will always help you
in classification and text summarization.
Many other aspects again,
dealing with data. And majorly, we use all these libraries
in machine learning,
artificial intelligence and LP natural
language processing and deep learning.
Also computer vision. This is the applications of Bear this particular library
will be used using PIP installed.
We are always installing all this library single handedly,
not in my library installation.
Every library will be installed
along with the execution output command stating
it has been installed.
It will show you it doesn't start now. You don't trust. How do you check it?
Verification of the installation
is very important because since we are working with machine,
sometimes
it might help you to have a better vision
when you verify.
If you don't verify if that installation
is crashed, you never know. It will affect your project
to better verify once you install it.
Very simple. In order to verify
as that you can open command, prompt
and type python double hyphen version. It will ensure
it returns the install python version.
What is this particular version you are working with next, verify
the installation of the libraries.

For that you have to just go open
python interactive shell type python
and then import every library
which you have already installed. If it is imported properly without
any error, then it does install properly.
Okay, so this is the overall development
environment setup idea which you have to have in
which you have to create in order
to do coding,
in order to create certain applications? All with the project.
Now we are in the command prompt in the article aspect.
We have known about various libraries
in Python that has numpy flask
stream led dart transformers. Let's install the same libraries
with the help of command prompt
first. If you could see it's not general, but
it's in my personal part. But it is inside right now
in my personal laptop.
The location of the python
is being fetched for that I have to.
You will see the command change directory based the location where you are.
Python has been installed
and then press enter.
When you do this, the command prompt goes
to this particular folder.
Let's start with the first bip install
numpy. Now, since I've already been working
with python a lot, many times for many projects
you will get the output.
Just wait and watch. I click on into it. Might take some time.
It will try to analyze what's happening,

what they're trying to install
and the requirement already
satisfy. Right? This is what you'll get. The output
that means number is already installed
in your particular system
because they're already working. And there is a warning is you could note
that if you want to upgrade the
particular library which you are using,
you can go for the version mentioned. I am currently using 20 1.2.3.
It is suggesting a paid for 20 4.1.2. Then what is the command for?
The same is also be mentioned here. You can use that command. We have now installed numpy
which is already existing.
It has given the message if in case it's a new installation
of library, how it will display.
Let's try with other libraries as that the install
flask I'm giving. So let's wait for the results
again. If I could see it states flask is already
present that dissatisfied again.
You have a warning. Got to the position. I almost installed all the libraries
but let me check
for the next one stream. Let
is a good see how it is downloading the stream
lit library. If you are trying to install the library,
which is not in the current
local system in your
this is how it will start loading you forget already existing.
This is the message
which you got for by end flask when you try to install stream lit
library
that is not present in your python. This is how it starts downloading
and takes 5 to 10 minutes

at least to complete the download
depending upon the system configuration.
Likewise,
you can install all the libraries required for you into your system.
Right. So I have given two examples. One how it will download the library
which is not in your system
if you already have downloaded
the library. How the message will pop up
that this requirement
is already satisfied
that it is already in START. Right? So this is how you import
your libraries in Python
Now in order to verify is your particular
library is installed or not?
Plus it will try to prompt you
that it is already existing. If it is not,
it will start downloading as mentioned.
Now, again, in order to verify that,
you have to go to Python interpreter.
So I'll click onto the same location type python here.
When you click enter,
it will go to the Python interface where you can execute the code.
Now what you do
is you try to import numpy.
When you try to give this particular instructor shown to the python Trump
inside the command prompt
which we have now, then it will try
to enter on import this numpy library
which is already existing
when you type the statement import numpy. If your number is present,
it will not throw up any event.

It will look just like this this indicates your number
is there in the python library folder.
This is how you verified the libraries. This is already installed before
using arrows of you mentioned in your code
as the it will throw up an error
if it is not installed before. Make sure you install the libraries,
then use it in your code.
This is so simple demonstration
of how you install and verify if the library is present in your python
with the help of command prompt.
Now let's talk about black
tagged with the app. We are integrating open.
Source. To our application. That is the main agenda of Let's
understand more about this application.
How do we work with this? And also look at the demo for this thing.
What is the basic setup we need is as simple as that
which is mentioned before.
Flask ChatGPT App
You have to have Python
installed in your system and all the libraries
mentioned to be installed in your system.
That's the basic ideology
for all the demonstration. We just got it in the session. Here are
the components
we need is plug for web framework. As I mentioned,
we use flask of Python library
for web framework and then open IGP API for generating
response is a simple logic.
We are taking API key of open API
putting that in your python code and then we are trying to execute this
same first lever.

Check how this particular code look like and what are the indeed
this step in our Google collab.
Note I'm not executing this
in Google color, I'm executing this in command prompt,
but for a better clarity,
I'm using the online coding platform
that is Google Collab in order to how good
interactive and bifurcation
between the text and the code. Right? Google Collab is a very good place
in order to work, in order
to have a good Python content on, let's
check out the code and understand
what does it do in order to create the app
using Flash library in Python
Now here we are on the Google collab. First step
we have to set up the environment.
As I told you, we will activate
the virtual environment of python.
What does this python M be? And V Let's understand one
by Python invokes
the python interpreter
which is already installed in your system. Then m b and the this option
always tells Python to the E
and V module as a script.
This module is used not to create virtual
and don't meant on that. That is why we try to this next again
you have v in me.
This is the name of the directory where
the virtual environment will be created and it is not mandatory
that you have to give the second we eat
and we as it is,

you can change this to be ABC D also, or you can also put it

as virtual environment.

It does not mandatory that you have to use

the same name, but with M you have to

use V and B does this mandatory

and the second B and B is optional. You can change name accordingly.

Naming convention can be changed

according to the requirements. Next,

after doing that setup, we create the Flask application. And as I told you,

we are using only two commands of Unix

since we are working in command. Prompt plus disk make directory

m get the odd that s we are creating a folder

with the help of command prompt.

That's it. There is nothing great that's happening. Creating a new folder. The folder name is GBP Chat app

and you are trying to change the path. Change directly

c be to that particular location.

And remember wherever you have installed your Python software, that folder is just

these things to be created.

We have to first navigate

to that particular python location. Then only we can create a new folder

instead of the execution

problem and partitions. After doing the folder creation, we will have to

create goal. Find first is a python code file

which I would like to name it best.

AP dot b by application dot. But again, this is not mandatory

you have to have a, b b,

you can name it accordingly,

but you have to remember what you have named While you're executing

this, you have to remember

the exact Python file

name, including cases in this case entity. Okay,

first we will import the flask elements. First is Flash library.

Next request

Jason ify and render template.

Then we will import the requests

and also import bank. These are the libraries will try to import

which is available

in Python to our particular code. We will try to use request

and render template

also the timing in order to have

the conversation between the system that is to be D which

we are trying to create and our questions.

Then we will initialize the flask

application. This is the last application

in initialization syntax

within right to give up open API key.

This is a secured key which you should not share

with anybody else, but as they can utilize

and you have to pay the bill for the same

but you keep the API keys very discreet.

This is a random API key. It's a sample API key or errors

you can just put in your code

into your API key here. This is far better.

Rather than giving people original API.

Once you put this API key, you will next define

the root of the home page.

Then it has to interact from. Obviously you cannot show this back

and go to that user.

You have to have a frontend,

you have a friend. Then that called Indexed Dark Patch.

DM But there comes the second first one that is having a dark break

or Python fight.

Next it has to be to the front

that does index dot hedged right

then render the particular hatch

DML template that is they will be using render template

library.

Okay. We have usage of

these libraries every that. Then we

will define the root for chat endpoint,

which the post request. Post request is nothing

but what message you put to the GPP

and what it has to respond back

and this thing will happen with the help of JSON,

then you get the response from GPT three.

And remember there's lots of GPP models which when you are using

you have have the knowledge about that.

You have 3.5 Turbo 16. Okay? You also have just GBP 3.5 tarball.

You have many kinds of model. The current model which we use,

you have to mention it here

and you also have to carry the input

given by the user from the frontend to the backend.

With the help of this message is the data should be transferred

from the front end backend.

What is the maximum limit of the response is 150 letters.

It's not words, it is very minimal. If you want to make it more obvious,

can make it 300 again.

It is according to the requirement

you are featuring four. Then you have to attempt

to get the response from API

which tries in case if it fails, now can still ward off it

at handling the code,

which not having error. Handling capacity is not over the code.
Simple as that. But if something goes wrong
first you have to let the system give you
the message that something is wrong,
not directly line to the error page. It should be interactive
and it should tell the user
whatever you have entered is wrong
or something has happened. What has happened?
This particular responses
should always be there. For example, I'm giving you possibilities.
It's not that every quarter will be doing
everything what they have to do right.
But still there are a few standard error
handling techniques. When you work, you have data score 200.
When this particular 200 status for can for not for error comes for 29 comes.
Right. How do you handle that? What is the error and
what is the particular response you give?
For example, if it is for 200,
we can return the message the user rather than going to a random
wrong page, you have to give a message.
An error occurred
while processing the response. Stop open API. That means if you are GPP
is not connected properly,
it is not able to give response. Then you have to just not push
that particular code to the page.
You have to send the error message. You have prepared this next
you have for 29.
Here we are trying to open
if fail status again, we are trying to re attempt
how many times back off retrying attempts
at the will try to do two time attempts

and then we will go first loop.

That means we are putting

this particular system into sleep that it is not able to start.

The more you work with that,

the more limit the exceeding will happen.

And for 29 also means that if you are particular open,

it is out of limits.

It is not having any limits left. It is actually that you have to buy a new

you have to put building again.

It will see

you have exceeded your current quota. Please check your plan on building things.

So that's how you have to try

to give error message to the user so that they understand

something is happening.

We have to go that bridge because nobody will go back in and debug

what is the error rate in the frontend.

It says we have to show what is happening. So this is sample

example of error handling.

Then if any other error comes more than this,

there are two errors which are listed.

If anything else comes up, you have to just give the status, go directly

or not for error or internet.

This connectivity error

anything might come. An error occurred while communicating with open

is a standard default error message.

You can send if you don't know what you have to do,

just put there is an error. Please decode

then comes to on the flash application. This is the main method which create for this code and the code

execution starts from the main method.

Here. Now comes the second part, the frontend which we had discussed already that

and what are the complete content of that?

We will have a quick overview. That is index start here. You Can see people who know it.

Steam will always know this

that you have had to have it. DML, Lang, English and Metacritic.

This that is always that you have style

for your particular page and you also have the body here

you have a checkbox box,

you have a text box,

you want a button, it's a simple thing. You will have to have a text box,

the user will put their input,

click on send a button

so that it will interact the to do with the display the tangibility message

you have to have a label or again

a text box. So that's how it will look.

It's a simple JavaScript

which is being used in order to have this interactivity

that's fetching the information

from the input and putting that to DPD

and taking the response from the DPD and pulling back on the front end

to view for use up.

So this is the simple fundamental function

that happens in this script section.

How do you run this? Just you have to be in the location

where you have created the folder that is DPD chat app.

Right? I'll just go back

and just give you a overview DPD chat app.

That is the location where

the command prompt should be pointing out. Then you can execute Python app

dot B right?

Then you do this,

you will be able to access a browser

when it is loading
in this particular address. What is this address by only 5000?
By not 4000? You My question, as you all know, it's DB 120 7.4.4.1 will always deal
with local hosting
when you do this local hosting, you have separate ports for every library
or every kind of execution you do.
5000 is the port number
which is allocated in every local system
for flask, library of Python, Any flask that framework code execution
you do it will launch on the chrome
with this particular address.
Hope you have a complete and detailed view
of how this particular application look.
A quick recap. You have to install python
and necessary libraries. Then you have to create a mean python
code fight that is APB dart.
Be right. Then try and say that is frontend
interface. You have to make it index targeted demo.
Again, this is your requirement. This is a common name which we keep.
That's why we used the same.
You can create a simple estimate interface
to interact with it's AGP. Then you can run the application
and check for the output.
Right. So now that we have understood
what is the code at the back end, frontend and every aspect.
Let's execute this code
and check for the output before going to the demonstration flag
with the app.
Let's understand the folder structure. I am here in the location where my python
has been in start simply worked on.

You could see many folders here. We are trying to create the chat app right according to our steps,

we've already done it so.

If you go to this particular folder, you can find two different elements.

One is templates are the one. This app,

what a template template is actually

the index file which we had discussed. I app is the main program.

Once you execute this, the back file will be created if the execution is successful

or not successful, it doesn't matter.

Once you run this through interpreter,

it will generate automatically. That's very disk present.

If you are executing for the first time,

this will not be death. Okay, Now let's open to the command

prompt and check out as we work with this. Here we are in the command prompt

and in the location that. Our

fight is. I'm right. No, I'm actually wrong. We are in the location where Python.

It's now

we have to enter to the folder created. What is the folder we created?

We have to choose directly

to that particular folder. GPP underscore check underscore app.

This is the folder name

which we generated. Right? Let me enter to that. See, now we are in that folder.

How do we execute the steps

which I mentioned? We have to type python and we have to mention the app

not be right or the name

which have given to the mean Python fight. You have to mention that and click on

and once you click enter

after you do this, after we verify this is how the output looks like

are read the right output screen.

No, it is just indicating

that it is executing. It is running the location.

This we have to go to. It should be. 120 7.4. 0.15 thousand. The port number.
Let's quickly hop on to that location
on our search engine. Any browser that you can use,
you can go to this particular port number
that hop on to that. Once you go click on
I should be the same already
that it has been launched across,
you will find that interface. What you have to do,
you have to communicate with that DPD.
So I'll press hi and click on Send button. It will say hello.
How can I ask you to do the next question,
which I ask How are you When you do this,
you send this.
I'm just a program so I don't have
any feelings, but thanks for asking. How can I help you? So this is how it is trying to interact
with the human being.
If you try to give which is not existing still,
your child is not trained to that level.
It is a normal basic model. I see. Where do you
I click on the send button. You have exceeded your current quota. Please
check your plan and billing details.
It did not drop this error. It's really if your limits are exceeded
that this then it will show this error.
It will try to do that error handling
which I have already mentioned. So for the third conversation itself,
how did we get this message?
You might be having this particular doubt. The thing is open API
API is not very much free to everything.
You only have access for $5 worth
of conversation that can happen.
API key that can generate

that is how you can convert. After it exceeds $5.

It will try to ask you to block and select the plan

and do the billing right.

The payment should be done for this thing. So this is just a simple example. You can enhance, can buy a paid version

and start building the products and help

the small scale business if you own any, in order to have a private chat board

so customers can interact without any

actual agent service required customer service,

you need not take it. You can use the box there on your website.

This is a simple idea. This is how the execution looks like. By now we have understood

how tag DPD why using Slack,

how we can execute what the code required

and how the outputs look like. Now that we have understood

and also saw the demonstration

how does tag DPD app work when you create with the help of flash library

using Python?

Now let's check out the next topic using the steam flask.

Flask Text-to-Image App

How do you use text to image application? Yeah, the simple idea is fixed.

Image generation involves creating images

for textual description using air models.

You will give a simple description Here

we are not focusing on description. Are trying to get up image for the world

which we

as I told got dog any animal or what

you want to fetch The significance

of this particular application

enhances creativity and design processes useful in various fields like advertising,

entertainment and virtual environment.

They Want to get an image? You can give a description

that it is sitting on a mac

or dog which is sitting on a bed. You can give certain description. You will get the images in certain

or sketched that image

drawing of the cow. So you can view a certain description. It will generate back the output for you.

How do you implement this particular text to image app? First, this will build a web application

that converts text description

into images. Again, if you want to build a web

framework, it is above blank.

Then you use open again. SDLC is this for the frontend

that is very much mandatory and basic?

What are the you want for this post? Is Python to be installed in your system?

Next you have to have a required library

that is like an open API and you have to have an API key

from openai.

This is the basic requirement

it should have in order to start off the development

of this application.

Now let's understand

what is the code for this particular app?

What does that purpose

and what are the use here Then? Later,

we will execute this of the idea.

Now let's quickly hop on to Google collect

understand more. Moreover, this application

here we are on the Google app.

Demo - Flask Text-to-Image App

First step. If your python is not in start

in Somalia by then if already exist ignore simple as that.

Again create virtual environment. We already had the description about each

and every element

of this particular code statement. Then we activate the virtual environment
by using this particular code here.
Next comes installation of FLASK,
an opening. It's very important to install the libraries,
which is necessary for your coding.
First place will be using the command
to install flask open it. It's a simple statement here.
The code line you can just execute same. Then we have to create the project
directory.
Again, nothing but the new folder
it is named. That's plus underscore text.
Underscore to underscore. If you want to put some of that name,
it's left for you. You have to go to the folder
which you have created.
Then only you can start creating your
python code file and animal code file.
First thing is main python application file, which is again
named as APB dot BE right.
It gives you a proper signification. It will log mixed with the previous one
because the folder is different.
So again,
we have to import the necessary libraries, then initializing the flask application.
You have to have an open API key. You can replace this. You are open
API key into your orginal open API key.
How Do we do that? Then you have routine which I still go.
That is index. It's demo in this HDMI file
you have all the designs
related frontend, the data
that has been fetched and will be routed
with the help of function call generate image as opposed to method.
We'll be using. Jason Push method means the response

which you from tangibility.

Right? Either it might be a user import

also act as post and also the response will be also forced to use open

API to generate

image based according to the prompt

which is being received. The size should be only this much and the

number of images at once should be one.

Only then the prompt which is given by

the user will be pushed to open API.

It will get a response and then that

particular image will be displayed.

If you want a detailed explanation of what does every line means,

I have it for you.

You can just read it. What are the different elements? The use write that we use right

next to you have to create this HD interface

that is the front. Then again, you have to a text box

a button and value

plenty of prompt

and then you will put that inside the DPD.

It will fetch the output

on the same screen. They do have to have a simple text box

and a button.

That's it. If you want to do more styling,

you can more welcome. You will see spines, you can do it.

This is the general basic setup

or the front end which you need. And you have a script.

Again, your function calls an image here. What happens to fetch the information,

the prompt from the user,

and then it will put

that particular prompt to open API. Once you get the response from open API,

it will push back the response

on the front

that this is the code for that same right. It's just an interaction code

between the frontend and back.

It will be JavaScript. This is about the HD file will be using. Again.

If you want to run this code,

you have to divide them AB Dhabi. Right? And you have be on this theme folder

where you have created at the start.

If you go somewhere, navigate to

some other location on your command prompt and if you try to give

it will not execute.

Let's have a quick recap. Your first efforts will be installing

Python next. The library is called an opening.

I laid out the creative task application

that this app be right. Then we will also try to include

the functionality of converting text

to image that whatever the text we have

given related images would be provided. So we have to root for that

and you have to have a simple HDMI

interface in order to have the connection

between the user and the system. Then you have to access and run the code,

the flash server, and again,

you have to go to the same location

that is the same IP address

which ends with the port number 5000.

So this will be your particular location

address. It will run there.

You can execute this. See.

Now let's quickly check. How does this work in that demo? Now, here we are in the Python location,

the local system we are going to execute. Plus next to image app.

If you try to go to that particular

folder, you could find the theme folder structure.
It is having a mean python code
and then you also have index in that simply right once it is
executed, back file has been created.
So that is right. They are here. This is the structure. Now what we have to do go to command prompt type this particular location,
try to execute with the help of Python
app that we recommend. So I am copying this location
completely going to command prompt
and changing the directory
to the copy location. Now we are in the folder block text
to image application
which has been generated. We straightaway try to execute this with the help of Python app dot B right command.
Once we do that we will click on and
this is there the flash server is been running,
it is active. Now we have to go to the location which is
mentioned right here with 5000 port.
Let's open to that particular location. Once you come to this location,
you could see
the basic standard design
that we have here. And it's our time
now to give certain description
regarding the image
and try to generate that image. I'll give just one line of description
you want.
You can give in the description
so that the deputy will give you. All right, perfect.
Required image as per the command. I try to give mountain
lit skylights. Okay, so
let's mention the color

also that is that will be good. So mountain lit green skylight.
This is my description of the image which I need I to generate.
Okay. This is how we the image from the GPP. Because given the light
which is in green color on the sky
and mountains are right here. So this is how the descriptions
will be taken care of.
The more precise description you give,
the more precise image you will get. Basin output.
So this is how the image application
will work with the help of open air in your python,
which also help to generate
the image just digital content creators or any kind of creative
people who work in that particular field. Having learned about them,
they iTunes and Lynch.
Let us know if you are dried. Any DNA. I don't see it. If so, what did you find most interesting?
Share your thoughts in the comments below. Up next, we level up our skills
with launching a powerful framework
for building advanced applications
will also launching for them to do
and even down to rank using link team ready
to turn your knowledge into action.
Let's get started. Let's begin
by Transforming Theory into practice.
Where
we'll be learning regarding Python 19. Let's quickly hop onto the agenda
and check out what else we cover in this.
First, we will talk about introduction to launching,
and then we will try to understand
what is the basic environment setup
we need in order. Work with applications
using long chain on Python

and obviously
with the integration of opening. Then we talk about core concepts.
We have to understand
followed by the components of land chain. At last
we will have a long chain case study
that will be understanding more about. How does this work in practicality?
Why wait? Let's quickly start with the session. Here we are in the first topic
introduction.
Introduction to Langchain
Do like team first. Let's understand what is like. Team Langton is a framework or the library
which has been designed to streamline
the development and deployment of
applications that utilize language models. So this is the exact definition of long
term,
why we always call it a library. Generally,
we used to do one kind of operation
that is installing the libraries. We use Web command, right? We also use the command to install
long chain in the Python environment.
That's why we consider that
in a normal way of talking. It is also our library, right?
So there's no doubt it's a library
but also they mentioned it is a framework. It provides a set of tools and components
to make the language model work in an efficient and stable,
scalable manner.
It is just like a playground where you build certain models
by using different kinds of elements, and particularly long chain
libraries are framework.
What are the key features of long chain? What does it include? It elm wrappers
when you see large language
models, wrappers or nothing. But we are having a small model
which is working as a prototype

in the 19 little library by itself you can write to use that really made one

or you can create new wrappers.

Then You have prompt and prompt template prompts are the basic building tool

which users will use

in order to communicate

with the ubiquity are important. It is English like statement.

It is just a requirement. You are putting it into our system. So when? So I knew this is how you work

charge you right?

You give the prompt.

What is your requirement in detailing? That's how you need. Then you will get the output accordingly.

So we are communicating prompt act as a language

between the greeting and the user.

As simple as that from templates. Yes. You also has templates. It will give suggestions just like

you could see suggestions and GPP.

Nowadays

what you are trying to ask the CBT or any other coding platform if you want

it will write to prompt the next code

or next word which you want to type. Generally it happens in Visual Studio

if you have observed,

so it is having a template already

or it will auto generate

if required. So that's how it works. Next, you have chains. What does this chain let's not go to long

chain aspect.

Let's talk about a normal chain. Chain is nothing

but interlinked elements, right?

It will link the elements

like this one to the menu link.

Then it is the chain. Likewise here, when you work a complete project in a module vice,

you want to link those modules

in order to make a complete whole project. So change
that, always act as linking goes together
and making it the whole pattern. So that's the use of chain. Next, embedding and vectors
tools obviously embedding,
if you to embed any other content from outside the world of long chain
Python or Openai,
you can use this embedding like. This does you have the content
which is already present in the
internal storage of long chain. So we also mentioned this to be a vector
stored, right?
We just address it that way. These are the key features of long chain.
Hope you had an overview
and understanding about the key features. Now let's understand
why do we need long chain?
Why LangChain?
We need long chain
in order to overcome certain challenges. Okay, what of this challenge is
why do work with Elam?
That is language models. Integration issues might come
up, difficulties in scalability.
I did solving for 20 people
now we wanted for 60 people are multiplies into 200, 300,
so on and so forth.
So by providing more than that or user
friendly framework, always, Langton
gives you an open window where you can
explain or demonstrate your ideas.
Eric admits. You can make a model and do the job next. It particularly specifies
it is having applications
all over the domains majorly. It will always try to reduce
the work of the human.

That's how it will give advanced capabilities with customer support,

intelligent automation

for content, recommendation system,

whatever the models, which will be your requirement,

it will try to give up

advanced level support

and processing capabilities. If you use 19 now, what are the use?

What are the applications? Where do we use long chain? It's used in every other domain.

You could name it, you have it,

but I've listed very few of them. First comes customer support.

When you talk about customer support,

what happens here? I'll give us an example for you.

If you use chat bot, which is built

with the help of long chain framework, it makes the customer support very smooth

without human intervention,

even though if the chalkboard

is not able to solve the issues, then you can intervene. The human 5 minutes of the time is saved,

Human resources saved.

It heads in time saving as that. Next, you have content generation. If you are a person who is a blogger who

works with content content writer, right.

You get you want some ideas. How do you have to elaborate content? How do you have to create a blog?

You can just give up from that. I need a blog on Python. It will give you a blog

you can just use that as your inspiration

and you can try to modify them

and put up into a requirement.

Then comes intelligent automation again. Automation. You think it's not only that you

how many other robots.

Also, if you want to build something,

a basic robot, you can use this framework

in order to create that

thinking ability by using language models,

which is available in flagship.

So intelligent automation

has that semantic search. If you are a person who is working

towards the domain of your search engines,

you want the ability to be proficient

in your domain.

You can use long chain

in order to have a semantic search. Also, you can use this for data

as the person who is purely working

with data contained creative values

or technical wise, you can use like cheap.

Then comes personalized recommendation. You want something to watch for yourself. You want to create your own

automated main integrator,

which having the ability

to give you this task which you have for the data

or any prompting you want in your system,

you can create

a personalized recommendation model. The help of long chain. These are the use

cases are applications of Langtang.

Now let's understand what is the software

or environmental requirements for lighting is a huge requirement

or it is simple requirement.

Let's understand them. No, we have Python in our system, right? So the basic washing of python as all

you need, nothing else

where you can import the long

chain library and you can walk across. So it is supporting three point,

seven and later versions of Python

interface. Next, what are the libraries

or dependencies you have for long chain?
It is very common you have python. TensorFlow transform was done by pandas
for data manipulation and analysis
and flask of fast API. So for deploying the web services
applications
TensorFlow for training and interface
and you also have transformers for pre-trained language models.
So these are the libraries
with all the dependencies you have. But it's not that
everything will be working.
Everything. Every library will be used. It's a broad structure library
where you can use development environment,
so you will develop this
in which interface. It is all about the Google
CoLab or Jupyter.
So these are the preferred IDE
for development. It is not restricted only for this. You can also have various ideas
in order to work with long chain,
but your ideas about Python
will all be supported 19 as well
because they both are integrated.
Now let's move on to development
environment step. Let's understand
what all we need already be hard work,
Development Environment Setup of LangChain
but let's see in detail
what is the command we use? How do we work trust? First thing with it, as
I told you, Python already
and 3.5 or later versions. What is the installer we use? It is always pipe
python package installer.
We use PIP command. Generally we talk it in that language. So install line.

How do you install? Just open your command prompt and run

the command called pip install language before that,

ensure that you are in the directory

where the python has been installed. After that pip install long

chain will work then verify installation.

Very simple. Go to python ID and type import 90.

Once you do that you will get no prompt

nothing, no commands and it will be a plain input.

Then it will be a plain output. That means it has been installed properly. If you want to check if it is installed,

then it will print this next statement.

You can give it for verification.

If you don't want to. The plain output

you can give a print command line chain is successfully in start.

It will be in the same. If it is not installed,

it will not go to the next line when it does not go to next line, it is

ensure that it is not installed properly.

These are the verification types you can try in order to know

whether it does install properly or not.

Now let's quickly

check this in the command prompt and understand

how this installation happens.

Demo on Library Installation

As I told you first,

I am navigating towards the Python location,

which is installing the local system.

Then I try PIP install long chain.

I'll type this and click on and once I click on Enter it will try to analyze.

Yes requirement already satisfied. That means I've already installed

this particular line chain.

That is right. It is not giving me any loading

kind of output.

It is telling everywhere that is

it has been installed not any long chain. Any packages related to

that is also installed

and it also suggests

there is version 24.1. Point two would you like to upgrade? You can use this command

and upgrade for now.

It is 20 1.2.3. Then it goes back to the Python location.

Now if you want to check, I suggest you click on Python.

It will go to Python interface. That is 3.10

which I'm using as I mention this.

Anything after 3.7,

this particular library will support what was my verification method, which

I mentioned import long chain. When you build this,

it will go to the next line.

Right? So if you want to confirm whether the line

has been imported properly,

you just have to do this. As mentioned, import 19. When you click on Enter,

you can also type a print statement.

Print chain is successfully installed

if this executes. Obviously

when it does come to the second line

that has statement,

it is already installed. If the import launching is not working,

it would have shown up error by now.

So this is how do you verify

19 installed properly or not. So we had an installation method.

If it is already installed, it will go to this kind of output

and to verify you can do this,

you can enter the python prompt

and then you can give the comments. Hope. This is clear.

We had a successful demonstration

and how it looks value install or you try verify with the help of import

command on the command prompt,

which is both Python

ID as well as the normal command prompt. We use both of them in order to verify

we use Python

in order to install,

we use normal command prompt. So that's about the development

environment setup.

Now here we are on launching code

concepts topic. First, we understand launching

LangChain Core Concepts

is built around the concept of modularity

because we use this library in order to lock in modules,

then make it the bigger project.

It will always allow users to easily work

with the language models by creating managing chain and agents,

which can maintain

certain state using memory. So what happens here? A core

concept for this includes

chains, agents and memory. Let's understand what are these chains?

What are these agents

and what are these memories? Customers try to understand

what this chain here chains are Sequences

of operations are components

that process inputs and outputs,

often using language models. So we try to have a sequence. Chain operation are the components,

which is inside

the language model of long chain. We to generate the output one by one.

What are the types of chains you can find? It's very easy, simple

as just like a simple statement
and sequential is you to go one
after the other. Only you have the particular rule
when it comes to conditional.
If case will comes in picture
if this particular condition been
executed, then only you can go further
or go to the next chain or next agent.
So this is how you can have
a conditional chain. You can put a condition for your chain,
you can make a chain
to mandatorily sequentially execute, or you can keep a simple chain
which integrates all small modules.
This is about the chains. Now let's understand what are agents? Agents are entities
that can make decisions
based on that inputs, interact with chain and utilize memory for maintaining split.
We want somebody in order to communicate
between the link. The module right chain will always act
as the link between that module
who will communicate from A module to B,
You have to have agent you want to have. A statement will come from a location
to be location or a module to be module.
So isn't this simple. It will try to establish communication
between the chains between
that small module that you have connected
with the help of chains, types of agents.
You have reactive agents,
proactive agents, and we did so reactive is just that
you will tell something.
It will give a reaction for an action. That is a reaction. That's how this agent work.
But our active agents

or it will always have an eye on what's happening, where It is

which module communicated with whom

the log history proactive

agents is always used for the log. His this session has like

when did the log in?

When did the log out? What happened? So what is the session time? So they are proactively behind the user

who is using

that particular project

or the content, whatever it is. So the active

agent is only when you walk them to react.

That is where they will be proactive

without any provoking

the actions you do with the system.

All know that that's

how proactive agents will look. Next. Are interactive agents just like activity,

like you interact with the system

with the help of branch, any other medium

or coding anything as such. So these are the types of agents you have.

Now let's talk about the memory. Memory in chain allows agents to maintain

state across state interactions.

It's not that you interact

and you keep quiet, but I have to have those data on off, for example,

that is also state.

Yes, no is also our state

enabling more contextual awareness

and coherent responses is always possible

with the help of memory. Start the memory.

How long the memory the name suggests

is just like an almond rom. So term is only for that

particular session particular thread.

After that it is that long term memory. It gets listed the words
the start of your project,
so it will have the log of it. When did this happen? What happened? So all of this is stored
in the long term memory
and you want to reuse that
to multiple aspects as a requirement. You can do it, but jog
the memory is just like you use it there.
You finish the work, you closet,
you are done. You cannot go back there
to find what has happened.
Right? This is one of the important core concept
which we use. But chains and agents, this is all about
19 core contents. Now let's know
more about long chain components.
LangChain Components
Long chain components involve a few of the major
aspects. The first one is prompt.
We all are familiar with what are prompt
nowadays because of time to right.
So prompts are essential for
guiding behavior of the language models. In a layman perspective,
it is an interaction
between the GPP
or the language model or the book and the use of use from a language
in order to communicate that
they define the instructions are questions
presented to them. So DPD will ask certain questions do that.
It will give an answer for us. It is having variety of options
in order to load your file, upload
that and involve a statement
in the prompt stating Analyze the updated Excel or any of the file

and me so and so output.
So it is having the ability to analyze. So we are evolving with language
models and very quick
designing prompt. We need to have certain clarity,
context and specificity.
It is very important. You cannot put the machine into confusion
with all the kind of
instructions
you give or the prompt to give. You have to have a clarity first,
what you want, what is your requirement?
Then give the context. This is the situation. I want a requirement like the one
so specificity I want exactly ABCD.
I don't want e object. So you have to give the second prompt
which is having all the three aspects.
One is clarity. When is context of the one is specificity. So this gives you a complete prompt
and also it is important
for you to make the machine understand
that's our basic job. So we have created.
That doesn't mean
it has to work by itself. We have to train in
order to work properly. Best
give prompt very simple, a wide complex instructions giving which is not able to process
by the human itself.
If you try to give it confuses model. The more you interact properly
with the model, the more it gets screened.
That is how it would look. So test and iterate. You try to continuously test the problem
to define them on the output you receive.
Try to give training to the model
use template if you already have template in order
to have a proper consistency and reusable.
If you are having a prompt ready

and it is used by multiple other people, it's easy.
If you have a template,
if your work is based on monotonous, you can just use a template
for the prompting as well. So
prompt will help the long chain or any language model don't train well.
Next you have models. 19 supports various models
each suited for different tasks.
What are the types of models
we have depicted comes first, right? It is all DPD now.
It has been a world a lot. It is having 16 K Garbo DVD 3.5.
You have many kinds of models
that are available Hugging face models. Example body two Again it is
being able is just an example.
You have custom models, integrated models for other
libraries are custom built models which are integrated in your website
or anything for personal use.
So these are the examples of models and long chain
always support different types of models.
It's not only is it the to what you are,
I'll only support DPD. I don't support the
but it is having open forum in order
to support the languages and the models. Now let's understand the core component
that is tools like Jean offers
various to enhance
the functionalities of language models. What are the kinds of tools we have?
We have fixed processing tools, data
augmentation tools, evaluation tools. So in text processing
you can find organizations Deming and
limit ization tools. In data augmentation,
you can use tools to generate synthetic data
for learning and also evaluation.

It is very important
metrics and evaluation tools to assess the model performance, to get the output, how
it is working, the results you call it.
So you have to have evaluation tools
as then it's not a you generate and keep it aside
and start using it.
What is it? Progress, How to improvise that
in order to know the drawbacks, In order to the flaws,
you have to evaluate in a different way.
It's just like when you do
the software development, you always do
without software testing, right So there are aspects of testing,
like with evaluations of metrics.
So long time
also support evaluation metrics in order to improvise model performance.
Next, you have data loaders. So the fourth one, data
loaders, facilitates
the loading and processing of data
for language models. I give a simple example
right in this video that we will try
to upload some file to tag
DPD will ask the tag DPD to analyze the updated Excel fight
or any other file which you give.
It will try to analyze and it will give you the output
and what you want to do with that. Have to tell that DPD So file
loading is very important,
but that data is a tool which we use
and you have two types of nodes.
Loader
load data from CSP files for Jason Loader. No data from Jason points in long chain.
I'm talking about in particular,

we have components called this. Then it supports these two kind of sites.

So we are

we are about the components of 19 Corp we are clear

but long chain component again

we have data loaders, we have tools,

we have models and we have products. So these are the major components.

You have to understand now what is next. We have 19 case study. You

LangChain Case Study

may they talking about a simple case study when you are trying

to generate your own story

with the help

of opening a python line in library or

and with the help of that API key

which you generate and explain you regarding the code

we use, the installation we have to make

and how do we do

everything in the Google collapse and we will also execute that in command.

Prompt Why Google collapsed. It's very easy

to speech on Google collapse.

We lose command prompt

and how do we install all the required libraries and how do we save the files?

Where do we get the files from? How we have to save the structure. Everything is explained.

So let's quickly hop on to the Google

collapse and understand. Now here we are on the Google collapse.

What we are trying to understand

what is the case study posting personalized story generate.

That means it will try to fix up an input from the user and try to generate

the story for the user.

This project will take inputs, might be the kind of phonemes, settings

and theme the story.

It will generate unique story

every time you try to communicate with Chadi easy points,

like why I'm stuck with 3.5.

That is for and for that coming right? But when you go to open

API, it's still at the 3.5 margin itself.

It is having 16 came 1105. That is some other code

that's going on is just a tarball 3.5

so many kinds of models but for now it is

3.5 in open API platform. Not talking about that today but or 4.0.

Okay. Don't get confused with that

but steps in order to create this project. Very simple

setup. Then we don't collect user inputs. Generate story using your model and

the state generated stood simple aspect.

But when we talk about that

installing of libraries. Yeah,

we have to install open a long chain.

It is related to opening

also related to python. So bip install opening

and matching both of these libraries

we have to install then in order

to collect the input from the user,

you have to have two different python

files here. So that is a differentiation.

Previous demonstration

we had only one python file, one file. We used to work with the code. Here you have two files.

What does this do? First, once this user input dot break,

that means they are trying to

welcome the user

and take the input from the user. Here you could see. Welcome to the story generator.

You have entered the main character's

name. Please enter that then into the setting

the story and into the theme of the story.

For example, Adventure, Mystery Order, anything at such an event,

the characters setting team to that

particular file called Story Generator. Don't be that you are taking input

with the help of one python file

and you are trying to that particular

collected input and the python file

that this story generates

from long chain import chain from fixed model, then user input. As I told you, user input file,

you have to take all the user input,

get the user inputs what is collected,

get the name settings and the this is collected.

You have to take this as an input

and import that into this particular story

generator by end, by

this is the function method. You can

just create certain story according to the input given

by the particular user.

Then you will have a prompt. You will have to work with the text model. As I told you, this is depicted 3.5 here

and here comes your open API key. You have to put your secret key here

and then you have to execute

this particular main block.

It will try to give you the story

generated here in the print statement denoted stories

so in a paragraph of a story

will be displayed for you. So this is how the main block

will get executed.

And these are the commands

and the codes are to be used at record

as having a self explanatory comment

that you can read and understand once again

if you don't follow it here.

Okay, This particular learning material

or code is always provided nobody.

You can go back rework on this again. Then what do you do

in order to display the story?

Have to execute the file. How do you execute story

generator dot B right here.

Right. Are we not using Python? We have the python, right. It should be python story

generator dot bay.

That's how it will execute. Simple as that. You will have a quick recap here. Post is setup

that as you are having python

in your system, you have two libraries

that is opening and launching. Then you have to have your own API key,

which is discrete.

You have to create main script

that is story generator and you also have to create the subscript

that s user inputs.

Then you have to. Do. The story functionality

that is open is deep.

You have to use that. Then you have to give them

the character name setting team, etc. It will develop the story

and if you execute

it will give back the story which is already developed,

as simple as that.

But this is very clear for you. Now let's see the demonstration

what is the output and how it will work.

We are trying to execute

personalized story generator, which we have already discussed, that

we are using the library code 19 year.

You have two different python fights. A story generator one. S use it in book. It is already explained.

User input is used to take the input
from the user and story generated is the main app.
You should not execute
python user underscore input not be right. You have things a good story underscore
generator start be right.
That's how you will get the output screen. This is kind of special execution
that every output is seen on the command
prompt. It says we need not navigate
between any other locations for output. What does this buy cache?
If you click on this folder
after you combine code, this is actually generated compiled
python file will be generated.
So that is right. It is here. Now let's quickly
hold on to our command prompt and try to execute this particular code
fight.
In order to do that. First,
what we need is we have to copy this location
that it is actually situated.
The folder your app. Now we are on the command prompt.
We are changing the directory
and pasting the location which we copy and clicking on and we are in the folder
called Personalized Storage generator.
What we have to do there,
try to execute python story underscore
generated dot be right once you click on enter.
This is how it starts executing. Welcome to the personalized story
generator and I'll type in name
of the main character as Ali's and it does ask
setting of the story there
it has to happen I can see Enchanted Forest.
I gave the location

visualization idea for the GPP. I click on Enter.
It should be a mystery run or adventure
run or horror run. Whatever you can mention that I mentioned
as adventure story.
I'll click on it to see the stories
generator in this form. You can read the pausing the screen, but
it will include the main character,
the setting of the story. Also, what kind of story?
What is the theme of the story? It try to give you the complete paragraph
which you can use it for your requirement. This is how our story generator will work
using long chain,
you can create much more applications. This is of one basic
example of this is clear for you.
As you know what is large language model will going to discuss
about limitation of large language model.
So why we will going to switch on drag. So before that we should understand
the limitations of elements.
Limitations of LLMs
So the first limitation
is computational resources to train the data and to deploy the model.
We need high cost then. So this during training the model
and deploying the model
using alum there,
we need high computational cost. So that is the first limitation of
our next is data dependency. So I am using vast amount of data
so they're actually highly dependent
on or heavily dependent on data. So after data dependency,
next is performance issue.
So I am having this limitation
because factually it provides sometimes
incorrect output which is referred

as hallucination after performance issue. Let's move to the next limitation
that is ethical and social consent.
So as you know, NLM
uses a very large amount of data, right?
So during that, that data might have
some sensitive information. Well, so we have to be very careful
about the privacy of the data.
And there we should be careful about the all the ethics for data,
what we are using
and all the legal things
related with data. So that limitation, what we have in
editor is regarding this ethics
for using the data and the data privacy
after that. Explainability
And last is generalization limits.
So these all are limitation
of large language models and because of these limitation
we move to that.
That will reduce that limitation
will be used to overcome this limitation.
Let us understand about a retrieval
augmented generation, which is right now what that we would hear
Introduction to RAG
Boston discuss about the definition. But then after we'll use
why we use this rag, the purpose of rag
and last will going to how it works. But before we start this, I tell you
the meaning of that retrieval retrieval
basically that odd is looking like a brain that search for all the information.
So what we can see here that art is basically work like a brain,
get that search all the information
then after this generation. But what is this generation

to that generation?
Basically, it's a creative part of that. Okay. So that use that is used to work on that,
the data what we have so through retrieval and then after the last part is knowledge
base after once we finish the definition
and the purpose and how it works,
we're going to understand the knowledge
base part is that the knowledge base is the last thing where we basically use
to collect all the information.
What we have gathered
after this retrieval and addition back.
Okay, So let's understand first
what is definition of correct. So right combines language generation
with information retrieval
to enhance next generation. Basically it makes the data more accurate.
Reliably. It uses external data sources
to provide more accurate, as I said
before, and a relevant response is next comes is purpose of rec.
Now the main goal is what Of course,
if we are working on data,
if we want train a data,
if we want to deploy the model. So basically we need the accuracy,
then we need the quality,
the right and reliability should be that. So these all three terms, what I said.
Is. They're indirect. So the purpose of rec over here
is the main goal to improve the quality
accuracy of language model output
because that issue
we are getting in a little, as I discussed
before, the limitation of loop.
So there is performance issue, right? So in performance issue, what I told you,
the output we were getting to learn,

sometimes it's factually incorrect
and but the output how little in them sure is like very confidently lln
provide that incorrect output
which referred as hallucination. So here that lag is used to overcome
through that limitation.
Here we can see that
to improve the quality and accuracy
and output, what we are getting
is more better and accurate and reliably.
They help models provide
updated information and contextually. Which. Responses.
So basically in brief
we can see the output. What we get through them is very accurate.
Very reliable and. The information, what we get through. Read at last very up to date.
Now the next point come is how rag books So here you can see the regulatory
relevant information
from a database or the internet
before generating the text. So database
we can use sequel or any other database
to retrieve the relevant information. These Retrieved information is
then used to produce more informed
response to get the better output and uploaded information
so that so that looks. Now the next you can see on the screen
is benefits and example of like the first will discuss about the benefit
benefit of right you can see that
the VEC improves the obviously accuracy
this is the best and the first mean benefit of fact
that it improves
the factual accuracy which we generally
don't get in sometimes in
of delegated text. It allows model to provide

answer based on the latent available data.

So whatever the available data we have, based on that,

it provide the correct output there.

Whatever

the information we have, based on that, it always try to provide the accurate

outcome there

and as well as uploaded information

after that, you can see that the use case of right

when a what is it common application.

You can see there we can use

the rig is in good said mode virtual assistant

and content creation tools

to add what you already know. A lot of said work

now we are using virtual assistant

which actually assist you virtually. If you ask anything, you will get

a response based on your query.

So in many application we can use direct and there

we actually get the proper accuracy.

Next is a lag is useful in. Any scenario where accurate

and current information is crucial.

After that, let's discuss about the example of so you can see on the screen

that I have given there

that Google search on window

model and open air models using retrieval.

So this is well-known examples of system where we are actually taking

the benefit of these

these done set the benchmark

for integrating retrieval with generation. So basically this system

you can see an opening I threw in opening

I we have a lot of models like you butanol there we are using these dragging all

because they are giving the the benchmark. What I said before that this system set

the benchmark because the output accuracy
we are getting to that
output is so far than because of track.
So this system actually said the benchmark
for integrating retrieval with generation.
So this Google says augmented models
and open air nowadays are using drag
just to get the proper outcome
in a very accurate manner.
So we're going to understand
more about read. So let's dive in to the read more.
So now the next concept is basic terminologies of like, let's understand
one by one
all the basic concepts
and terminology which we use in
first is retrieval component. Know what it is. So the retrieval component,
as you can see on the screen, it searches
RAG Basics Concepts and Terminology
for the relevant information
and it actually looks like a brain whose search
for all the information, right?
So the retrieval component. So just for the information and from there
it says from latent base
or some other source like Internet, then after this component and shows
the model has access to the up to date. So this retrieval component,
when it serves the information, the key
point here is that whatever information it
search that should be uploaded.
So it always focus on that also. And after that, why
it always focus on updates because if the
information what it will search for

basically if it is up to date, then based on those information
it will gain the accurate final outcome.
So here it should be first uploaded
the contextually relevant data.
Right? It should not be something as if suppose
we are searching for a data, it's
completely relevant
for what we are searching. Next is.
Generation component. Now what it is
the generation component to use
retrieved data means the data what we have
so through retrieval component.
So it uses that data
and why use this to generate some response
based on that uploaded information. This end show the generated text
is both coherent and factually accurate. So generation is used to get
the information
from the retrieval component
and then output it is stored on it and try to generate the accurate
and reliable data.
Or we can see the outcome. Now, the next you can see here is
very important, which is knowledge base.
These three are very important
parts of the So knowledge
base is basically what the knowledge base
is, the source of the information. From there you are getting that
information in retrieving component.
So here you can see on the screen
that the knowledge base is the source of information
used by the retrieval community.
As I said before, it can database

or other sources. From there, the retrieval component
is actually retrieving the info
so it can be any source can be a website
or it can be any database like sequel,
basically the structure collection,
all for the information. From there
we can there the proper up to date
and data is basically comes under knowledge base.
So from where we are taking what from where we are retrieving
the data is knowledge base.
Then once we retrieve the information,
the important and useful information, the uploaded information
then will go to generate component
where this component is. Start using that information
to generate our response,
which will be very accurate and reliable. So now the next one
is contextual understanding.
So you can see on the screen
that this actually allows the model to the user's equity
accurately.
This has in drawing
the most relevant information by generating a response
that basically I'm understand
what is the context behind
the Retrieved information.
It all allow it try to interpret
the user's query accurately.
So based on that query
only will going to generate the response. And then what we want
that will come in the outcome right.
So if it'll understand that properly, then it will generate the accurate

and reliable response.

So this is contextual understanding. Next countries relevance scoring. Now what it is?

Well, it was a scoring ranks. The Retrieved information based on its

importance to the query. This ensures that the most pertinent

information

is used in the generation process and basically it gives you the ability

that the information which you are using

is the most relevant information

to generate that response.

After that comes integration of retrieval and generation.

Basically the retrieval and generation

when it will get more is basically the final part.

The seamless, integral version

of retrieval and generation

component is crucial for VEC.

This integration

allows what for real time information, the literacy rate and generation.

So basically the last part of this,

which is called the merging part,

basically when you get the proper

retrieval, the process is done,

then when you will going to generate that seamless process.

It comes under the integration

of the driven generation. Now The even going to understand

the benefits of using ride

with Element

was why we are using the bag with limbs. So as we have already talked about

the limitation of adding to to overcome

from that the limitation we were going to use because it extended

more beneficial in terms

of for hallucination intelligence. We have disclosed about that

also and a lot of other limitation.

What actually have in our lives

while using a lens so to our control that we are using here.

So let our mist in step

by step. Everything. So let's discuss all the benefits of.

Using. A drug with NLM. So first is enhanced accuracy,

as you know that the while

using LLN, we actually have a limitation

that the performance accuracy is not up to the mark.

So here it actually work on it and how it work

it always try to gather the information

or to retrieve the information

which is up to date. And if the information is up to date of

you will get the correct response at last.

So it ensured the generated text

was fully be reliable and gone.

Correct. It should not be like

we are working on today and we are getting the outcome

of the studies like

so it always try to fit

the updated information. If we are looking to try to fair

to the addict and updated information,

what do we have to do? And based on that one will get the response.

Obviously that will be more reliable

and correct. Next is contextual relevance

by retrieving relevant information. You can see here

the rag shows that responses

are contextually appropriate. It should be really stable.

Then this leads to more meaningful

and relevant interaction with the use it. It should be more relatable.

It should be more relevant. So then only user can get their response

proper.
Did it try to go down the relevant info
and it tried to fetch
only the relatable data. Next come is handling
diverse equity
that allows to handle a wide range of topics,
but accessing a vast knowledge base,
it enhances the model's ability to answer questions beyond pre-trained knowledge.
So we always train the knowledge
base, right? So beyond that limit, we try to give
the accurate and proper answer
based on the user's query. So that's
why the third is handling diverse queries.
Next is reduced hallucination,
which is a very important point over here,
which makes it a better version, makes it why we are using nowadays, why
we are more focusing on read with limb. We call it help
minimizing the hallucination
where more to generate incorrect
or non-essential information. Because while using language models,
most of the time we face
this problem of hallucination
needs the model confidently provides, provide or give you
the incorrect information right. And based on that,
when we get the response, of course
it will be not accurate,
it will be incorrect.
So it tries minimize that problem
to try to minimize that hallucination.
You can see here
the retrieval of accurate data supports more credible

and coherent next generation.
So at large, when we generate the response
that will be more correct,
accurate, reliable and relevant
and updated as well. Next one is after that scalability.
Now you can see on the screen that. There can. Scale to include the new information
without need of the training.
The model means
you need to retrain the model. You need to train the model again
and again
so it can scale to include
new information. Basically, if you are working on something
for and if you are getting more involved
later, the update of this,
it works on uploaded information. So it include that uploaded
and new that new information
it can include but no need to. Again work on training the model.
So without retraining the model we can
scale to include the new information
it makes it easy to maintain
and update the model's knowledge base.
So the database knowledge base
is what means when you get the data.
When you stored the data in the database,
read the database automatically, get a so it helps or it makes the process easy
by maintaining
and updating the model's knowledge base. Now the next one is
what is the Dell application?
So drag is beneficial
for various applications, which is you can see here that customers upload
content creation and virtual assistant.

So it has lot of application
where we can use this VEC its ability to provide a precise
and the relevant information enhance
the effectiveness of these application
because these all are those application
which actually need the accuracy more.
So the point
that is why we are using this drag and this application
because they are actually
dependent on the accuracy more and it actually because of this drag
these application
provide more effective outcome. So these all are the benefits
of using that with large
language model novel going to understand
by using this simple example. So now next is key components in agility.
Key Components of RAG - Retrieval and Generation
What are the key components of it? Do it. The first one is the driver component. The overview.
The retrieval component is responsible
for finding relevant information
to support the query response.
So if I use it, ask a query. It is start working. Okay, it start trying to find the relevant
and related information
based on the query to provide
the accurate response through generation.
It not only use the pre
trained knowledge base, it used the internet
also to get the relevant information
based on the user's query. So text assist external data sources
database or the internet
for this information. This to get the uploaded info. After that comes knowledge base.
This is again the very important component here. Did knowledge base is what it include

the collection
of all the information related well information collection
of document collection of articles,
what we are getting through other sources. So it serves as the primary source
of information for the retrieval process
after that efficient search mechanism. So here you can see
the retrieval component fullest,
understand the query context, do
so it effectively. It involves what it involves
passing the query to identify key dumps
and concept that if you ask a query it actually make this search process
more efficient.
How it tried to get or try to find the important details out of your query,
tried to work on it
more to give the correct
and accurate response.
Now the next one is information
retrieval algorithm.
Now what it is, as you can see over here,
that algorithm like beam to file
or be if I if I use to find
and rank relevant document,
relevant document, as you already know,
that those who are more related with query of the user.
So this algorithm is called document based
on their relevance to the query terms.
And based on that,
only we can say that we are we are getting the accurate response
because more relatable things
only can show you
the final accurate response. Next is a relevance scoring.

Now, what it is a do document that is going to determine
the importance to the query.
So it prioritized
based on the ranking of the info. And so the most important
will be on the fourth rank
then so on, so high scoring document are prioritized
for generating the response.
So it will move to the generation estate. Then after that, again, you can see here
efficiency mechanism,
it will advance certain out
so this advanced mechanism to it ensure the quick and accurate retrieval
of the input of the information.
This efficiency is very important
or crucial. We can see for real time applications
like DEADBOLT and Virtual Assistant
next game is key components of generation. Let's talk about first
the generation component of the review.
So the generation component overview
is what it uses the Retrieved information,
the information, what we had agreed
to retrieval it, use that to create the response and the response
will be very relevant to the query.
It integrates the Retrieved data in the meaningful
and contextually relevant reply.
So whatever information that retrieval process retrieve through
different different sources, it actually integrates
those data based on the ranking.
After you can see here it is
contextual integration. Now what it is,
the model integrate context
from the query and retrieved information

and then after this and shows

what the response is not only accurate

but very appropriate. Also as per the user query.

Then after next comes language model. Know what it is?

The generation component is actually used. There is advanced language model IGP

three Djibouti for these model

have been pre train on lost data

all ready to generate human leg fixed.

So if you ask a query,

obviously you only can understand if it will come as a response

in like text.

Right

then only you will be able to understand what you are getting as a response. You cannot understand

in the form of encoded text.

So basically is indeed the human like text to make it more easy

to understand for a user.

And it works on a large amount data

because these models

have been already trained

pre train on vast amount of data.

Next come coherence and fluency know what it is.

The generated text is designed

to be coherent and fluent

and resembling natural human conversation.

This quality is critical

for user satisfaction and engagement. So here,

if you have noticed what I have explained

before that suppose if you have asked a query read with the model.

So there if you want to refine it. Right, Because as a response what you are

getting, you are not satisfied with that.

It's not like it's not accurate. But the point is
you want something more for as a response.
So Here you can do the conversation in a very natural human conversation
in that manner.
You can do the conversation
and you can know refine the response in the order
based on your query.
And based on that, again,
you will get the more accurate outcome. So this quality is critical or important.
What for us means for user satisfaction
and engagement. Next is dynamic content generation.
The generation process is dynamic
and allowing what the model to create of diverse
and adaptive response.
It can handle various type of queries From that
A question do conversational prompt.
Conversational means you start asking
again and again based on the response,
what you are getting through the model
and based on your refinement of equity,
it start giving
you more better version of the response. So this is a very important
component of generation.
Now, the next is response optimization. Now what it is, the model continuously optimized,
the response based on the feedback.
A new data feedback means we want to give something
more based on the response.
That feedback loop will work over here. So based on the feedback
and I suppose something come up, right?
So based on that new data, it always
optimize the response in a better way.

This adaptive learning
actually improves the accuracy and the relevance of future responses.
So this adaptive learning
not only improve the accuracy, but actually it improves the relevance
of the future responses also because it continuously optimize the output,
or we consider response
based on the updated information
and based on the feedback. What we share
or what we give to the generated response.
So now, as we have discussed,
as you spoke about the benefits of using the bag
with aluminum, let's talk about
Workflow and Applications of RAG
the effectiveness
by using a simple example. But so far as you can see here,
good example.
So here you consider a user asking
what are the latest advancement in
a standard language
model might provide a general answer based on its training data.
Because the training model
based on that only generally
the language model provides
the concept retrieval process is the next the right model
for us to retrieve the recent article
research paper on Advancement. So this is the example of this up
to date information
like rag use always uses,
always the updated information. So this is the part that this initial,
then the response
includes the latest information available. It's only will work on the train data,

but if you have some uploaded information
regarding your query
so it will try to fetch that without
the training done data. Then comes generation process. You can see here
that using the Retrieved data,
the modern dendrites are detailed
and update response. There we go as follows.
The Retrieved data is uploaded already. The obviously once,
once the generation processes start,
so it will give you the updated response
and in a detailed manner.
This response is more accurate and
informative compared to standard models.
And the reason because it
not only look on the train data, but if you have something
new related with the query,
what you are asking for it, work on that and then will give you the final response.
So the accuracy is more here
and more informative. Now the next is comparative effectiveness.
What it is a standard
model might say is improving rapidly with the new technology in contrast,
and that model could see
a recent advancement
in AI include development in quantum computing ethical framework as reported
in the latest research paper from 2024. So you can see the difference
that the standard model
how will get the response to do that. And if we are talking about right then how
and what response
we could get through this. Next comes user satisfaction.
Know what it is. Obviously, the name itself is explaining
the whole thing that the detailed

and current response provided by lag
enhances user satisfaction.
You can see in the above example
that the reply, the response, what we are getting through standard
and the read which one is more satisfied.
So if I see the outcome through,
the standard model and if I see the outcome
what we are having right not to rag model,
the second one is more satisfactory. We are understanding properly what are the
recent advancement in the right.
So this satisfaction should be there for the users and user receive precise
and relevant information, correct
information,
reliable information, uploaded information which improves the overall experience.
So that should be there
it should be there for a user.
Because why
we are the working of online language model, why we are working more on
this to get the satisfied experience
from the user. Now next comes practical implication
of what it is.
As you can see on the screen, this example
highlight how I can be more effective in real world application,
how drag is so helpful
in terms of better user experience
and demonstrate the practical benefits
of integrating retrieval
with the generation for
interaction. Now the next part of workflow of how drive works,
so basic workflow of regularly
going to understand the the step by step

we're going to discuss about
the first one is user
query will work on the user query.
The first step in the process starts
when a user inputs a query
or a prompt into the system.
This can be a question
this can be a request for information
means if I want to know something about
anything, I suppose if I to know
about the latest advancement in so I can ask as a request
for that information what I am providing
or any text input requiring a response after this user query
when we start with our query.
Next is the trivial state that they start doing its work.
So the model searches for the relevant
information based on our query from a predefined knowledge base
or the internet.
It's not only depend on the predefined
knowledge base it try to fetch it, try to retrieve
the information through the internet
as well, try to collect
all that relatable, relevant information.
And then once it retrieves the most relevant information, articles,
data, whatever, based on the query, we move to the next step,
and the next step is ranking
Retrieved information. Now, what it is,
we have already discussed in a brief
that the retrieval information
is the rank, the based on its relevance
and to the query

how much that Retrieved information
is related to the query, The user query
based on giving those information
and the most relevant information
will be prioritized for the next step,
for the next stage. Okay.
So here you can see that
the recorded information is ranked based on its relevance
and importance to the query.
How it is important,
how it is useful, how much it is useful for the query, what use it is asking for.
This a step rich ensures
that the most important data should be prioritized
for the next stage
because it will not prioritize
if we're not going to rank the most prominent data, most relevant
data, then the accuracy might vary.
So here the step actually is very important
to have
because it actually do the ranking thing
ranking process, basically trying to prioritize the most important info
from the retrieved information or not. The next is after that generation estate.
So what then Mission Estates
do we have already discussed it in brief that model uses the top,
then retrieved information
to generate the response for the user. Here you can see it involved
what it involves creating common
and contextually appropriate text
that addresses the user's query.
Very and very clearly. Basically, it looks on the ranking.
We do information. Why? Because based on the top

prioritized information, which is what

we do, is more relatable to the query. Okay. So based on those top rank information

it is start generating the response

and start

giving the detailed and accurate response, which is more relatable

and relevant to the query.

So this called Generation Estate. Next is response and delivery. Now what it is, the generated response.

What we have discussed in the last slide

that the response so what we based on

what we retrieved it

through the ranking level state,

when we generate the response, that response is delivered

back to the user.

If we have logged on your query,

we have retrieved information, then we have prioritized the most relevant

information based on your query.

We have generated

the race response in a detailed manner. Now it is time to deliver it

back to the user.

It is time to send the final response to the user based on his or her query.

The response aims to be accurate, relevant

and informative

means the in a detailed manner

and handling user. So we have already talked about

we have already discussed

in between the Standard model

and the rag model. You can understand now the difference

between both that how we get

the outcome response from the standard

and how we get to the drag model. So here because of this that the response,

what we get at the end
will be more satisfactory
next is feedback loop. Here you can see that user
feedback can be used to refine and improve
the retrieval in generation process,
and we can see this also in sometimes. So maybe we can get the proper
refined or some thing. That model also can miss it,
try to give it accurate response.
But sometime suppose
what if we miss anything? So here this feedback loop. Do what?
Based the generated the response,
then it'll get delivered to the user,
then use it to get the response
at last user feedback Here
basically one user to give the improved feedback.
Look where we can improve,
where we can define it more. It's like if you give a prompt, right,
if you are getting some outcome
based on your product and if you want something more, it's
not sufficient for you.
So that information,
what you are getting as a response. So you will again ask back, okay,
you are going to ask back,
are you going to give the prompt back? I want more on this.
So that feedback is working here like a refinement on your equity.
So what it improves it improve the reply
well in generation process more
and this loop help continuously
enhancing the model's performance. You will get better
version of the final response.
So this is the feedback loop

after this one going to understand the key
of four that means for retrieval and for generation.
So let's discuss
now about the various applications. The first one is customer support,
how we use drag in customer support that enhances customer support
by providing accurate
and timely responses to the user 22nd user
Real time data Retrieval
To Answers Devil's question
on Improving Customer Satisfaction. Next application is Virtual Assistant.
Virtual Assistant is like what CV
and these applications
uses drag to deliver the relevant
information quickly.
So whenever you ask any question
or whenever you see the CV and Alexa, you get the response
very quickly and correct also. So these all because of implementing drag
to that the virtual assistant,
and it enables them to handle a wide of tasks from setting reminders
to answering complex question.
So it actually uses there as in to set various
different different does.
So that is not only helpful to get
the relevant information or uploaded
information, it is helpful in other task also for especially a virtual assistant.
Next come is content creation know how it can be helpful
in creating a content
so it adds in generating high quality content for blogs
articles ensure social media.
If you ask a query based on your information, what you will provide that
I want in this manner.

I want in that manner. So based on your information,
in when we do quality, it will try to you the correct
and accurate and detailed response
that so that can be helpful
to create your blog or article
or any other social media post. It retrieves the relevant information
or data
to ensure the content is accurate
and uploaded. Next application is medical diagnosis
How it can be helpful in terms of medical. So in healthcare Meg Assistant Doctor
By retrieving and summarizing
the latest medical research, they support accurate
and informed decision making in patient.
So suppose if you are not aware about the latest advancement
in medical research,
obviously it will not so helpful for the in terms of patient care.
Right. So if you have the complete info
about the latest research,
what is going in medical or in healthcare? So that will be helpful for a doctor
to take care of patient
based on the latest
in advancement research. What we have in the medical field.
After that,
you can see here that it is helpful educational to how it is helpful
in education as we use
in education to provide a student
with precise and detailed explanation
it has in creating personalized
learning experience by retrieving delivered to study material
based on the child or based the student.
We can provide the relevant study material

so they can go through that
and it will be more accurate
and up to date. So there also the right is very helpful.
Next comes legal research. So basically legal professionals
use this drug
to quickly find and summarize
the pertinent legal documents.
This is speeds of the research process
and improve the accuracy of legal advice. So you know that if we work
with the legal documents,
we actually have to go through
with a lot of documents just to read the document
actually takes a lot of time.
So it gives you
this summarize form of those document. And based on that summary, you can work
in a more effective manner that research
or you can give the more effective advice
there quickly. So that's what that is very helpful
in legal research as well.
Now the next one is
the next summarization. Now, before we understand
how it is helpful
to really understand
what is tech summarization. So text summarization is what you give
an iPad. This one an example in a very simple way. I am explaining that
if you are giving a very Linda
and really text heavy paragraph. Right. And if just want a brief about that
good summarize
that complete
but in a very precise a very brief manner so you just one
basically text summarization do it

focus more on the key words
out of that complete better and it is very important
for you to understand
so text invasion is what it is
a process of condensing a long document
into a shorter version
while retaining all the key information.
Greg enhances
this by retrieving the most relevant part of the text
to include in the summary.
So what basically begs do actually read the complete or large document?
Okay, of that
it tries to pick the important keywords. Okay. Based on those important keywords,
it tried to summarize
in a very short evolution. So that will actually help
to reduce the time
of reading the whole large document. And it's obviously save your time.
So this is called text summarization. Now next is automatic summarization
to that power
tool can automatically articles, reports and books. So it is very helpful
in text summarization.
And by using this tool, it's
actually give summarized version
of any report, article
or book in a very quick overview.
This saves time, as I said before,
and Therefore, for users who need quick insight from the text,
then just not to waste their time, we can summarize the large amount of info
in a very short evolution.
Next is improving comprehension. Summarize text Help the readers understand
the main point

without reading the entire document. What is the meaning of this? Basically that the entire document,

if you read it, will take time for you

to first, and once you will read the entire document, then after

only you can come on some conclusion.

So then you will get the summarized text

means When you will get the main point out of that entire document,

you can work on it in a very quick manner

and then you can come on the conclusion

also very fast. So this is particularly useful in academic

and professional setting

where time is limited. After that, keeping summaries relevant.

So by retrieving up to date information,

you can see on the screen that what is written that Dragon shows the summary

include the latest development.

It always try to work

on the updated information. This is very important for those things,

especially in

and research where the up to date

information is must very information is constantly evolving

means it keeps changing, right?

So that it is very important. So next after Will going to discuss

customizable summaries what it is user

can customize the level of detail

in this summary according to their need. If I just want to suppose

if I am giving a large document

and I just want to understand

the main point, if I am a beginner level

and the document is not only for beginner,

it is for the beginner, for the profession

nil or basically we can say for advanced level also.

But I just want a summary
based on as I am a beginner, this one day of summary
based on that book.
So it will actually customize
that again based on our meet
that can generate brief overview or more detailed
summary based on our preference.
So if I am a professional
and what the document I shared,
I want a detailed summary of that
so I can customize in a way,
if I'm at beginner level, I just need easy summarized version
so I can customize in that way.
So it is very helpful in customizable somebody after that integration
with other apps,
it can easily integrate with other apps. How you can see on the screen
that somebody can be integrated
into other applications that just email
client and document editor.
So if this flexibility, then it will be very helpful
for us in our work.
So this enhances productivity by,
providing quick insights directly within the tool
use it already use.
So no need to work again
separately because it can easily integrate with the tool
on which we are already working.
So that will actually
will make our job easy.
And this flexibility
should be there by using any to it.

Right? So drag is having that quality
or we can that benefit.
What we have to drag is that it
can easily integrate with other apps. Now next is advanced
question on settings system.
How drag is helpful over here us. We and missing in the introductory part
of what is advanced question
on setting system advanced Q system
utilize sophisticated algorithm to understand and answer complex query.
When we ask any question,
we should get the proportions. So in a very simple manner, I can see
what is question on setting system.
If you ask very complex query,
you should get to the proper answers so that the advanced users
don't understand that
complex squinty and give you answer. This system go beyond simple
keyword matching to accurate
and contextually relevant response. After that integration
of retrieval and generation,
it combines retrieval mechanism, gather relevant information from database uses
generation model to synthesize coherent and accurate from the Retrieved data.
So integration of retrieval
and invasion is what when you retrieve the relevant information
based on user query,
it didn't read the accurate response and the up to date response
and the relevant information
that is the part of integrating retrieval
and integration. But after this, a role of element.
Q Is system
now what is the role of added system? Large language model activity for enhanced
queuing system by providing

different context awareness, Ponce's
alum can handle a wide range of topic and generate detailed answer
based on vast amount of training data.
Because if you talk about DB2
for is trained by using a large and vast amount of data.
So if you give any query, if you ask any query using a prompt,
you will get some response
through that model for. So how?
Because it can handle wide
or different range of topic and generate detailed answer
based on that train.
They don't what we have we did the data
is a large amount of lost amount of data.
Next is real time data access. Now the meaning of real time
detects is is advanced
Q assistant then access and retrieve
real time data. Ensuring that answers add up to the real
time is what rate
whatever you are asking
it should be uploaded on real time basis. So whatever query you will going to ask
based on that, this will actually retrieve
the information and then rate the response
which will be uploaded.
This capability is very important
for application required. Current information, especially
for example for news channels, right? So it's a new and financial data.
That financial data, of course,
should be uploaded on any stock market. The data should be uploaded.
So for those application
it is very important after comes personalized response you is system
not only used for real time data access,

but can provide personalized answers
based on use it history and preferences,
personalize and as you then experience by making interaction
more relevant and engaging,
let's discuss about application across
domains widely used in customer support. We are already about virtual assistant
healthcare and education. These systems help automate response,
reduce the workload of the user and improve the efficiency across
different different sectors.
So this is all about various applications
in drag. We have discussed
about various application like how we are
using in customer
support, Virtual Assistant and many more after this
been going to understand more about
Red that the hallucination which is very crucial
or important to understand.
So let's discuss hallucination
in next step. Now we're going to understand
the most crucial part,
which is hallucination, indirect, how that will be helpful
and how does it help
this hallucination in end. But before that will understand how this
in that in explicit what is hallucination?
Hallucinations in RAG
Hallucination occurs
when the model didn't read the text, which is factually incorrect.
Again, this response may seem plausible,
but are not based on the accurate information. So in short,
we can see that when the model
then read out would generate the response. But it is not correct. It is incorrect.

So because of that,
we obviously lose the accuracy there and that is known as hallucination.
It actually this issue actually we have more in language model.
So next is because of how
this might happen why we get this issue.
So hallucination can occur due
to the limitation of training the data or the model inability
to retrieve the relevant information.
So in that case, this hallucination
or so they often occur
if the model attempt to answer
the question beyond its training scope. So if you remember, I
have discussed this, that rag is actually not only work on the train data,
but if the information is getting uploaded
right, it try to retrieve that uploaded information
also without the training the model.
So this thing
if is not there during the modeling,
this hallucination can occur after this example of hallucination
and allele
might generate a false historical event
or incorrect scientific fact.
For instance,
stating that Albert Einstein was born in 1900 is a clear hallucination.
So if the information in what we are
getting as a responds very confidently,
little provide you the information,
but it might be possible
can be wrong or can be incorrect.
So this causes the hallucination. Now we are going to understand
the impact of this hallucination on use.

It tells nation
can lead to miss information. Basically it mislead you
from the correct or accurate information. Deduce reliability of the system.
Because if you are getting incorrect information,
if you're getting an accurate response.
So obviously the reliability
on an air system automatically will get dropped down.
User may lost trust in the technology
if they frequently encounter
the inaccurate information then
so that trust what we generally have on a
that will get the proper
and accurate response automatically will get dropped off then after detection
challenges detecting hallucination
can be challenging as the generated text
or what appear coherent plausible. It requires verification again
reliable source to identify inaccuracy.
So if you get to the inaccurate,
the response, obviously you have to work on it again
and again to get the correct response.
And at last, also, maybe we're not sure
that what we are getting at last as a final response is correct
or is incorrect.
So that telling this we have to face
if we are having this elimination
in the millennium
after that need for addressing it, mitigating hallucination is crucial
and important for maintaining
the credibility
and usefulness of VA system and shooting accuracy and hands.
User trust

and the overall effectiveness of it not only enhances the user trust,
but if we get the accurate response,
then only we can proceed further
for the next step. So everything will get to stop
if we are getting the inaccurate response
or the response
which is not relatable or relevant. Again with the query
what we are asking for.
So here that need for addressing
it means that ensuring accuracy
enhances user trust also and the overall
effectiveness of all elements. But now how does rogue help
mitigate hallucination?
How would this helpful to overcome
this problem? So first is integration of natively
in this you can see on the screen
that red mitigate hallucination by integrating the retrieval component
that access accurate information
that could always retrieve the relevant and accurate
and updated information
to try to integrate that this ensure
that the generated response
are based on
and will be based on the uploaded info. After that enhance
the accuracy automatically.
It enhance because if you are having
the opportunity in for accurate info
so the accuracy automatically will get
and has an improved. So by retrieving relevant document or data
that provide oh factual
basis of generating text, This reduces the likelihood of the module

predicting that accurate
or fabricated information. Next is contextual relevance. Now what does it mean?
That right always ensures
that the Retrieved information portion should be very contextually relevant and
an native and accurate to the user query.
It should get match with the query. What we are asking. So this has the model,
then read the responses
that are both accurate and appropriate after that real time data access lag
not only look on the data, which is then we have in a knowledge base
pre defined knowledge base.
It tried to fight the
to date current information as to read system
can access real time information,
which is very important
to point out what we have in red,
that it can access real time information
so that today because of the reason it keeps model
knowledge gathering.
So basically if we get the uploaded information to accuracy
automatically will get maintained.
This is particularly useful
for answering question about recent even recent event or rapidly changing
feeds like news channel in news channel. We can not though like
keep the old or historical information
only. We have to be very updated there
and in financial market
of like a stock market,
that detail should always be uploaded
in those type of fields. This is very important. Have next is feedback mechanism.
User feedback on the accuracy of response
can be used to improve the retrieval

and innovation process.
Continuous learning from the feedback
help define the system
and reduce hallucination over time.
So if you continue,
learn from the feedback. If you continue our work
based on the feedback
and refine your response automatically,
it will reduce the hallucination issue.
Next come continuous improvement and the system are designed to improve
continuous lives
by incorporating new data uploaded data and refining the re retrieval algorithm.
So if you are getting a response
and if you use that feedback loop, so on that it to start working on
refining the retrieval algorithm.
And if it start refining that,
if it continues to improve on the information,
so you will get the accurate outcome,
at least
this ongoing enhancement for reduces
the occurrence of transmission
so that so that help
mitigate hallucination. After that we'll going to on done
this track how it is helpful in NLM
Steps to implement RAG with LangChain
and what is the process what all are
the steps that should be covered while using rag in
we're going to discuss next.
So now before we implement drag
using Lantern, we're going to understand
the steps to this rag we'd like to post.

The step would be install required
libraries. You can see here that begin
by installing lines in another necessary
libraries using the app, use
the command pipe, install lantern transformer F8 developers
and then CPU.
It actually helpful
because if you're going to use you have to install all the required
libraries first, then only
you will be able to do more for them. Then the next step is up.
The environment import required libraries
and set up the environment for Langdon and initialize necessary
configuration.
What configuration you required
for the trivial generation component that you need to be taken care of.
Next is compare data. You have to load the data and you have to prepare your dataset
to create a knowledge base,
because knowledge base is what basically
used to store the information from various bit is database A-Z
and from various resources
they have to load and pre process
your dataset to create a knowledge base. From there
you will be able to get the response
through generation process,
ensure the data in format suitable for retrieval
such as document or text based messages.
After the next step is initialize
retrieval component, you have to initialize integral component
by because based on the query
you start searching the relevant
information system That dual component

using language and building function

configure re retrieval setting to define

how the data will be accessed and rank

and then the configure generation

component after retrieval of course,

the generation component will be there.

So you have to initialize that component using a Pre-Trained language

model, IGP two, three or four set up.

The model would then read the text

based on retrieved information. So then you will set this.

Then only you will be able to get the right response

based on the Retrieved information.

What you have you then integrate

with during generation? Combine both component to create the final

system, ensure seamless integration

so that the data inform detect

generation process that best and refine.

This is the last step

where you can see that best. The Vex is done with sample queries. We're going to

then after give the query sample query

and just to get this already that whether it work correctly,

whether it's working correctly or not,

they find it driven under

written parameters based on the result for optimal performance.

Now we're going to understand a simple example to demonstrate

this complete the process.

So first step would be install

92nd would be in both libraries,

third would be load and pre process data. Both would be initialize

retrieval component. Then after configure

the generation component

after that create read Jane and at last
best to the system
by giving the sample query. Now let's understand
the future of read as we have understood
how to implement drag in long chain, we're going to understand
what are the future of Rack.
So upcoming trends and innovation in regs are integration with real time data.
Future X system
will increasingly integrate with the real time data because now
the things are to keep changing, right?
So now the data, what we are working on again and again,
we have to refine it, right?
So if the real time data
is getting change and getting up to date, so this rack system
will increasingly integrate with those,
this will enhance the relevance
and timeliness of generated response, making them more useful for dynamic
application.
Next is improved retrieval algorithm. Now what it is, advances in retrieval
algorithm will allow for faster
and more accurate information,
and if the retrieval algorithm
will get improved more
so obviously that advanced part of the read one
version of that will give you better
result at last. So enhanced the algorithm. Making it more advanced will improve
the quality of the retrieved data
that will be more then up to date and means that update will be there,
but the accuracy will get enhanced.
That multimodal that system. Now what it is
combining text with other data types

like images, audio and video,
This will enable more contextually rich response
expanding the application of
it will be more helpful in various applications
where we are using.
Then, after enhanced security
and privacy, innovation will focus on improving data security
and user privacy in the next system.
Because if we use a vast amount of data, we should take care of the
the data could be sensitive and there we have to be very careful.
So this enhanced and privacy to what it actually secured
data handling practices
and it will be critical as this is
don't handle sensitive information there after scalability improvement
rack system will become more scalable,
handling larger dataset and more complex
queries. This will enable there
using and more demanding application.
Right now
we have discussed few application, some where it is not in use in in that level.
If we improve the scalability we can broad or we can use this
racks in stone
in more demanding application as well. Now the next part is we're going to talk
about
that potential impact
and advancement of advancement. And is it forced right will drive
significant advancement may be Acer
pushing the boundaries of what policy
what is possible with a research will explore new architectures
and new matter new
different

even though these improve the retrieval and generation after that
and how the decision making.
If business and organization
will use this right they they can enhance the decision
making process also access to accurate
and uploaded information will get
the accurate and up to date information so that will improve the strategic
planning and operation.
So basically overall, it will be helpful
in making a decision correctly and accurately increased accessibility
to information.
That system will make more accessible
to a broader audience.
We can use in no different domain where
we're right now if we are not to be able
to use this right there. Also, you can use the rag to get
the accurate response properly. Now user will benefit from accurate
and relevant information
tailored to them
need regardless of their background. After the revolutionizing
customer service, it will lead
to higher customer satisfaction
and more efficient support system.
Improved healthcare solution in medical. Also,
it will be more helpful and beneficial.
It will not only work on the in the latest research of medical,
but it will give more up to date.
Our What will be the future of that after the research
the private research of our information?
What generally provide you
right often that what can be there in the future

that also you can get easily
so this will improve diagnostic
the accuracy and patient care, leveraging
the latest medical advancements.
They will not only give it up
to date medical information and research and it will give you what comes
that also will be there
after the economic growth. So this is also an impact of
the widespread adoption of lag
will contribute to economic growth,
how business will become more efficient
and innovative. It actually save your time right?
So there obviously the efficiency
will get increase innovation obviously
will get increase to lead to new market
opportunities and job creation. So now will going to understand
how to use rag in why a hands on for us.
We are going to board some necessary
laboratories. So what they are and what is the purpose of the libraries
will understand a step by step.
So the first is will going to import
the first library, which is
being stored 19 community.
What is the use of this? It is used for building the various
application using language models.
Now going to give other library
another one which is C transformer.
So baby install C Transformers.
So basically it focuses on
efficient usage. Okay. Next we're going to import what? B IP
install transformer. So these are the hugging phase libraries

of for model like board DPD. Next, the next library,
which we'll going to
give here is now we're going to give
another one under the library, which is also very important use in this
and that is basically f8i double s
c b you. Okay so basically
this library's F machine,
it is used for efficient similarity, search and clustering
the dense vectors will get after
this will going to give another
one we just baby install
sentence for much.
So this library is for sentence and body. Okay. So next you can see on the screen that
piping star lang Jane hugging face. What is the use of this library?
Basically this used to be Lang Jean
all we say language model applications.
Okay. So these all are very important libraries
which we have to install
while applying the Reagan
while performing the hands on part.
So as you can see on the screen,
we have given all the important libraries
that needs to be install
before giving the code. Once it will get install,
then we're going to use the
libraries from these all. Okay, So next
we're going to import the libraries
and we're going to load this model.
So basically you can see that few things are downloading over here
and you are already satisfied.
Okay, So once it will get finished,

we'll move to the next quote. So now all the required libraries
that needs to be installed
before we start decoding,
it's already done over here. As you can see it, the install.
And if it is already there, then you can see the message required
all in this episode.
Now we'll going to give the libraries
and we're going to load the model.
As you can see,
I have given the code so from Langton Community, what we are importing C
transformer from transformer.
What we are importing the organizer
or the model Y, we're going to use it
that we're going to understand
step by step by using these libraries
during our coding.
Okay. So here you can see that
we have even the libraries and we have given all the required models
also which are going to use for further
modeling then and run this code. It will take time because step by step
it will input all the required
libraries over here.
Now next step is what will going to load the them
using C transformer?
Hands-on RAG in detail
How is it going to do so? For that we have a code
that L going to show you over here.
So now you can see what here on the screen
that we have created a variable l
11 which we are using C transformers.
We are using a model that is the blog now model seven Big Iman

and the type of this model Islam.
What is the meaning of this
code is basically we are loading the model by using C transformer,
so let's run this code
and it's done. You can see that
it it's fetching the details
and the model
which we have used from there. Once it will get done. Then we move to the next code.
So now you can see that we have used over here the
model by using C Transformer library.
Now the next task is to load the tokenize. The organizer and model for embeddings.
How to do that. First we have to create a variable
with the name tokenize it,
and in that we're to use word or auto tokenize. So from pre
train in that we're going to
type sentence transformer. What is the use of this?
I'll let you know.
So first in the organizer
we are using the organizer from that we are giving from reading sentence
transformer
and after that we're going to use all meaning l
l6 we to know what it is. This should be fun.
So it loads of pre-trained tokenize
tokenized. As you can see from pre-trained. So it actually load the reading,
tokenize it
from there from mini
alum and six v2 model. Okay, so this is the meaning of this
as what we are doing here.
We are loading the token nicer
and the model for what I'm building
so that we have done.

Now let's run this code. Now the next part
over here, you can see after loading
this we're going to load the mini model and Vivan going to that.
We're going to do this
for generating the embeddings. So first we have loaded the token nicer.
Now we're going to load what the model. So the next goal you can see on the screen
is loading the model for embedding. So once we load the model for embedding,
we have loaded, then
now that tokenize it and model
both body and building, but next step is what they going
to define them building text.
For that we're going to create a function. So first I'll give the code over here.
Define ember
text,
and in that in that user defined function,
we're going to give some input. So for that
we're going to create a variable input
token. Either module, use text,
return, then Now the next code over here
is after giving this,
we're going to use the third dot no grid function
and in that we're going to use what the output listed in a state.
So basically it is what we're going
to understand once all the code over here. So that is output output equal to model
within bracket input
embeddings what we're going to use this output.
But last
he didn't state that mean dimension.
We are using equal to one
and then CPU
not number

and at last written
this code is so is basically we are creating
a lot of user defined and it's
converting the text text into what
they are embedding using the model.
This model we have used over here
is, is mini L model. So by using that model
we are converting the text into Into what?
Into the embeddings. So once we'll run this code, then we can
use this function in not further process.
So let's understand what we're going to do
next. We're going to get the embedding
dimension, how to get them adding
dimensions that we're going to understand
by giving a code. So now you can see over here
that we have given step by step
all the code
first have given the building them in how to get this
by giving this code over here.
If we want to initialize the index
for the omitting that we have to create a variable
in which we use this library
and in that we have to take
the next bonded and building dimensions. Now, after that we're, going to store
the document in memory book store
that we'll get to store by using
this function will go in the store. Next is create
an index to a box write in mapping.
So in Western broadcast already we are creating a dictionary over here
in which everything will get stored.
Okay. Then after will going to create

a vector store in which we are calling
F8 Developers Library and in that
we are calling all the variable
whichever we have created.
Okay we're going to give in this function. And once run, this next process would be
what the last code basically
is doing over here, that vector storage
is actually creating phases.
So vector store and it manage
and so is the document embedding.
Now the next to prepare your document, how to prepare document for
that will going to look good.
So the next step
is to prepare your document. And so as you can see,
that I have given few documents pro
and in the document
I'm giving you statements, right? So when I ask at last a query
on this document,
I'll get automatically the onset of that. So let's run this first
and then we'll move to this step.
For now, what will be the step you can see on the screen
that I have given that a code that is text
equal to double dot page
underscore content for duck and document. For now, I'll give you a brief about that
that we are omitting the document
and adding them to the vector store. But the meaning of the first code
you can see on the screen
that post code over here you can see
text equal to dog dot page content. What is this what basically it is
doing over here, it actually
extracting the text

from where it is actually extracting the text text content from each document
it is extracting in the document list.
Okay. So the first is extracting
the text content. Next, you can see that over here
I have given a code for I am omitting
enumerate and buildings and in that
I'm using a loop basically over here. So it is doing what it is actually adding
each building to the FSS index.
And you can see here that I have given that index a store ID equal to doc
I page content.
What it is basically mean
what is the meaning of this particular code
is it's actually mapping the index entry
and there to the corresponding document
content.
Okay. So the meaning of this
we have already understood let's run this code.
So now the next one
is defining a simple retrieval. This code basically do what you can see.
I have given our user defined function
over here the define symbol to work within a bracket.
I'm giving a query as an input. So basically this function retrieves
the most relevant query.
Okay. So to retrieve the most relevant query, we are creating this function over here.
Let's run this code and you can see that how we have given over here that query
and is going to end with next query.
And then we giving that the
I call to index dot search. So these all are basically for
what does they do
by comparing X
and building to those in FSS index. So by comparing that from the first index,

we can get the most relevant query. Okay. Next is will going forward

and that will give the next code.

Now the next is creating our drag chain. For that we are calling a class

and with the name simple retrieval queue,

we so the meaning of this class, it

initializes a queue system post point section,

as I can see on the screen, that it is

initializing a QR system with what

we call a language model and retrieval.

Okay,

so here we are calling the language model. Also the idea of giving the code

for driver also and by using

this class is initializing up to a system after once we run this code.

Let me explain you the second part. We are defining a function over here. Then self query and context equal

to self to the auto retrieval.

Within bracket

we are giving query over here. So this method is doing what is actually

retrieves the context and for a query.

So whatever the query will going to give. So for that, actually retrieving

that context for a query

and generate an answer. So if a class, when you give a query

by using this, it will generate

the answer based on the query

what you are going to ask. Okay,

now then going to do what the next step.

So the next step over here

is what we're going to ask a question using the red model

as my query over here is what is Lancton.

So based on the documentation,

based on the function, which I created, based on the class

which I have called over here
to create the rank chain,
I will going to give this query over here.
What is length chain
and the answer from there it will
fetch Q under is called chain. But then within that the question
what will lost?
So Q A chain is basically what there
you can see here that we are creating collection and it is useful basically
to retrieve the config for the query.
So once we'll run this code, you will get the answer
that what is in over here.
So as you can see, after leaving
the question of you're asking the question using red model
and how we are getting the answer
from the right model,
then you can see what is purpose length. General is giving the complete output
based on the query
what we have asked of would him
so that so we can use the right little. And I want you now to go and check
all these codes applied manually and this
then how it works. We have explored the world of general
going what it is, its benefit and its potential to change the future.
You have learned about generative models, the ethical issues, babies and explore
popular laws,
language models like CBT, BLOCK 3.5,
Sonnet and Gemini.
We also showed you how to create an app for Android and introduce
some thought generative A.I.
tools from crafting effective prompts. Working on project Vicente
Pitti, Python and GitHub.

Go pilot. You are now prepared to innovate
and build. We finished the advanced topics like
launching and drag, enabling and enabling you to develop
cutting edge applications.
If you are. Eager to take your skills
to the next level, subscribe to our channel for more videos.
Head to by like going
to stay updated on new content, keep building, stay curious
and let your creativity guide you.
- Generated with https://kome.ai