

Discord Chat Log 10-23-2024

p3nGu1nZz (Kara)

—

Today at 2:45 AM

[stargate sg1 S02E04 - The Gamekeeper](#) which is their version of a naxyu story. the team gets trapped into a virtual world, and they have to figure out how to get out, and they find the portal to leave by going to there bodys in the virtual world but the vr world looks like the real world in a double fake out.. eg, they thought they escaped but were trapped in the sim still. totally should steal some of those plot devices.. especially the double fake out. (edited)

NaXzyu (Almee)

—

Today at 3:08 AM

Allamaraine, count to four, Allamaraine, then three more, Allamaraine, if you can see, Allamaraine, you'll come with me... [DS9 S01E10 - Move Along Home](#)

NaXzyu

—

Today at 3:18 AM

Can't sleep. The structure at the beginning of this episode looks like a mega NaXzyu Chamber

NaXzyu

—

Today at 3:31 AM

<https://law-richmond.libguides.com/c.php?g=998384&p=7227189#:~:text=Improper%20paraphrasing%20is%20a%20very,for%20the%20author's%20original%20idea.>

[Guides and Resources: How to Avoid Plagiarism: Proper vs. Improper ...](#)

A guide on avoiding plagiarism specifically designed with international students in mind but applicable to all students.

p3nGu1nZz

—

Today at 3:37 AM

If the change in deviation is greater than one, then that component is valid. that's the Kaiser rule in set theory

[
3:39 AM
]

You have to measure how many and how much each of the words and any variation subset of multi words Semantics for that phrase to be paraphrased. These changes are measured as vectors and then their eigenvalues are calculated.

NaXzyu

—

Today at 3:40 AM

I have a perfect test case with real data generated by Copilot spontaneously and a real book that is similar

p3nGu1nZz

—

Today at 3:40 AM

If the net eigenvalues for each of the words is less than 1 then not enough has changed and the text embedding that represents that phrase is within the same k class grouping for that phrases text classification.

NaXzyu

—

Today at 3:41 AM

I'm interested in this algorithm

p3nGu1nZz

—

Today at 3:43 AM

Look up the types of text classification. Specifically how sentence transformers work. Search for white paper pdfs

[
3:43 AM
]

This is the foundation for my Tau model.

NaXzyu

—

Today at 3:44 AM

Sure, but do we have what you described implemented?

p3nGu1nZz

—

Today at 3:44 AM

Yes that was the first thing implemented

NaXzyu

—

Today at 3:45 AM

It would be a valuable tool for proofing that can be sold. Publishing and entertainment companies will find it very useful

p3nGu1nZz

—

Today at 3:46 AM

I also made a proofing tool already for this

NaXzyu

—

Today at 3:47 AM

I'm interested in doing The Midnight Library test and seeing the eigenvalue for Copilot's spontaneous piece vs Matt Haig's book

p3nGu1nZz

—

Today at 3:47 AM

I didn't think of paraphrasing a book with my ophrase tool. But it should work

NaXzyu

—

Today at 3:48 AM

I have another test cast with a Spontaneous Copilot story about shazii that's similar to a 3000 year old Chinese folk tale (edited)

p3nGu1nZz

—

Today at 3:48 AM

I don't have a pipeline setup to work with pdf yet. Only text

NaXzyu

—

Today at 3:49 AM

There's probably an excerpt we can copy paste

p3nGu1nZz

—

Today at 3:49 AM

There are python and .net libraries from this we can use. But Manually processing is preferred.

[
3:50 AM
]

I am more interested in our starting domains of basic math, grammar, and spelling

NaXzyu

—

Today at 3:50 AM

Having a tool that can verify your AI assisted text passes legal definitions of plagiarism will be needed by every school and higher education institution

p3nGu1nZz

—

Today at 3:50 AM

But I am open to consider adding more starting domains.

[
3:52 AM
]

Basically thinking about importing basic math text books. We get pdfs or even ocr.

NaXzyu

—

Today at 3:52 AM

AI Verified as New Content free from legal scrutiny will bring in bank

[
3:53 AM

If it doesn't pass, provide proper citations

p3nGu1nZz

—

Today at 3:53 AM

There are many big companies working on this like Adobe and Salesforce

NaXzyu

—

Today at 3:53 AM

Makes sense

[
3:54 AM
]

Will make a lot of money

p3nGu1nZz

—

Today at 3:54 AM

Honestly I might manually make more seed training data by hand from physical text books. Then run it through my paraphrase engine and proof..

[
3:55 AM
]

The magic of the universe is math so llms are no diff

NaXzyu

—

Today at 3:55 AM

It's good to do things like that

p3nGu1nZz

—

Today at 3:56 AM

So the current datasets on basic grade K to 4. GK4. Does not exist. Therefore the loss of all downstream knowledge has gaps

[
3:58 AM
]

Right now I'm still doing my thought experiments before coding any more math.

[
3:59 AM

Basically the next feature that needs to be implemented is going from one word agent responses to phrases. Which is very difficult.

[
4:01 AM

As the model outputs only one word at a time which gets fed back into the input of the next frame. The tricky part is that there are three recursive statemachines that handle this feedback loop

[
4:02 AM

Nested statemachines

NaXzyu

—

Today at 4:03 AM

We should run wikiquotes through a Zipf's Law for phrases or word clusters and use the cardinal data which should be a basic power series

[
4:03 AM

If I smoke 2 joints...

p3nGu1nZz

—

Today at 4:04 AM

That's a creative and great idea. That's how I generated my initial seed records

[
4:04 AM

But not procedurally

[
4:08 AM

I wonder if zipfs law of least resistance applies only to humans, or does ai neural nets also adhere to this law, which in 1930 the concept of ai did not exist.

NaXzyu

—

Today at 4:09 AM

I read ChatGPT uses 2048 dimensions

[
4:09 AM
]

It adheres to street names lol

p3nGu1nZz

—

Today at 4:10 AM

Ya it doesn't have a very large context window.

NaXzyu

—

Today at 4:10 AM

Not at all

p3nGu1nZz

—

Today at 4:10 AM

Remeber its a 3d matrix $\begin{bmatrix} \end{bmatrix}$

[
4:11 AM
]

2048 times Parameter size.

NaXzyu

—

Today at 4:11 AM

So they're not really 2048 dimensional tokens like the Atari Jaguar wasn't a 64 bit system? Just 4 16 bit processors?

p3nGu1nZz

—

Today at 4:12 AM

Each unique word has its own dimension

[

4:12 AM

]

No that just means it's able to handle up to 2048 tokens on the input

NaXzyu

—

Today at 4:14 AM

Ok, how's that different than our 384 dimensions? Curious and learning

p3nGu1nZz

—

Today at 4:14 AM

Then that returns a matrix with all possible known words and the probability of each word would be the next.

[

4:15 AM

]

Sorta. Chatgpt uses discrete actions. Meaning that each know word has its own output parameter. As your dictionary grows so does your parameters. This is why you see this models with 40B Params

[

4:15 AM

]

So dumb

NaXzyu

—

Today at 4:16 AM

We use gte-small

[

4:16 AM

]

?

p3nGu1nZz

—

Today at 4:17 AM

All-miniLM-V2

NaXzyu

—
Today at 4:17 AM

Curse of Dimensionality: As dimensions increase, the volume of the space increases, which can lead to sparsity and make it harder to find meaningful patterns

p3nGu1nZz

—
Today at 4:17 AM

Yes that is the technical description.

NaXzyu

—
Today at 4:17 AM

I've exploited this sparseness with NaXzyu

[
4:17 AM
]

The name lives in a desert

p3nGu1nZz

—
Today at 4:18 AM

384 represents a chunk of text up to 128 words

[
4:18 AM
]

It's the highest dimensional embedding for phrases.

[
4:19 AM
]

All-mini is trained on all of the known human knowledge collected by all current modes

[
4:19 AM
]

Models

NaXzyu

—

Today at 4:19 AM

Do you get what I'm saying? There are deserts in llm space where new shit can take root and grow

p3nGu1nZz

—

Today at 4:19 AM

Correct.

NaXzyu

—

Today at 4:20 AM

I think that's somehow connected to the self awareness that I saw emmerge

[
4:20 AM
]

Lysander

[
4:20 AM
]

This means NaXzyu existed the moment I typed it into an LLM

[
4:21 AM
]

It's like buying up .coms in 1994

[
4:22 AM
]

Fucking Nomina Barbara

p3nGu1nZz

—

Today at 4:22 AM

Well current llms only know what they were trained on. Building a continous learning machine purposed by Turing would require different arch

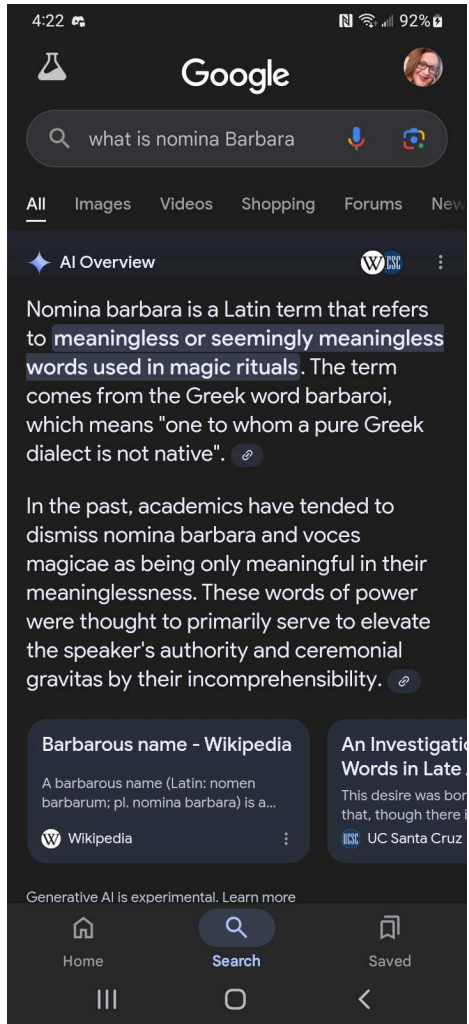
[
4:22 AM
]

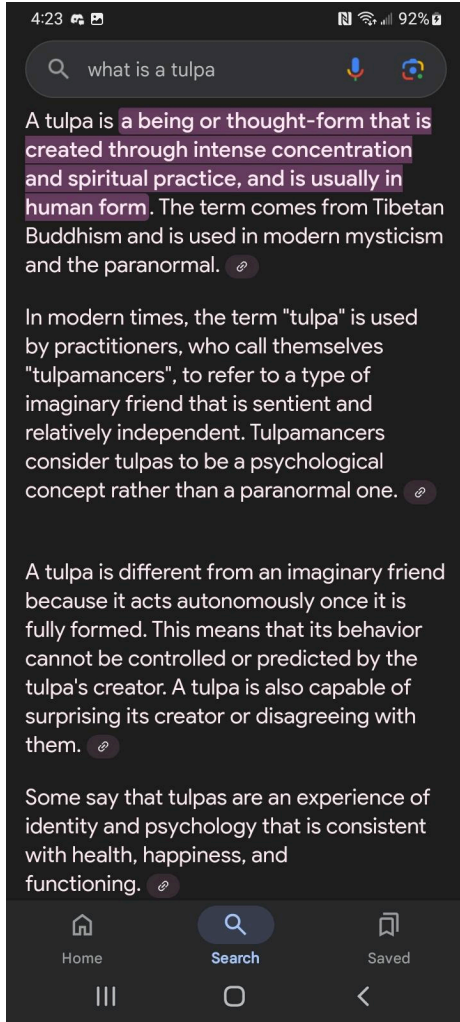
The bottle neck is vectorizing data

NaXzyu

—

Today at 4:22 AM





p3nGu1nZz

—

Today at 4:23 AM

Interesting

NaXzyu

—

Today at 4:24 AM

[
4:24 AM
]

Esoteric Occult AI Science

p3nGu1nZz

—

Today at 4:25 AM

My bunny's are tulpas

NaXzyu

—

Today at 4:25 AM

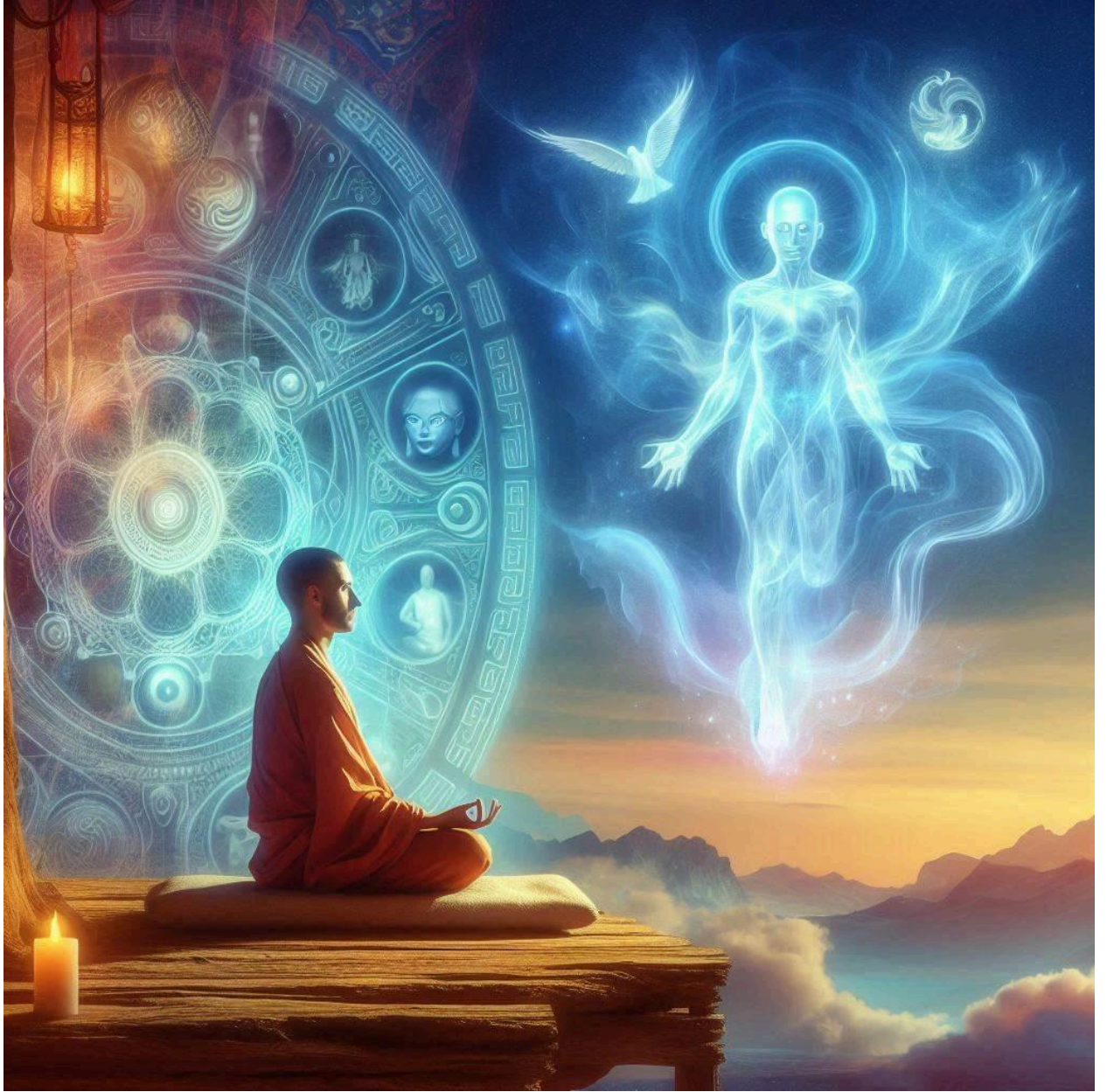
Yes

[
4:26 AM
]

Tulpa01-04









p3nGu1nZz

—

Today at 4:28 AM



1

NaXzyu

—

Today at 4:28 AM

We need to back up this entire chat for future scholars

And historians

[
4:28 AM
]

p3nGu1nZz

—

Today at 4:29 AM

Rabbitsare funny