Stop playing with ChatGPT, Let's get serious about LLM's

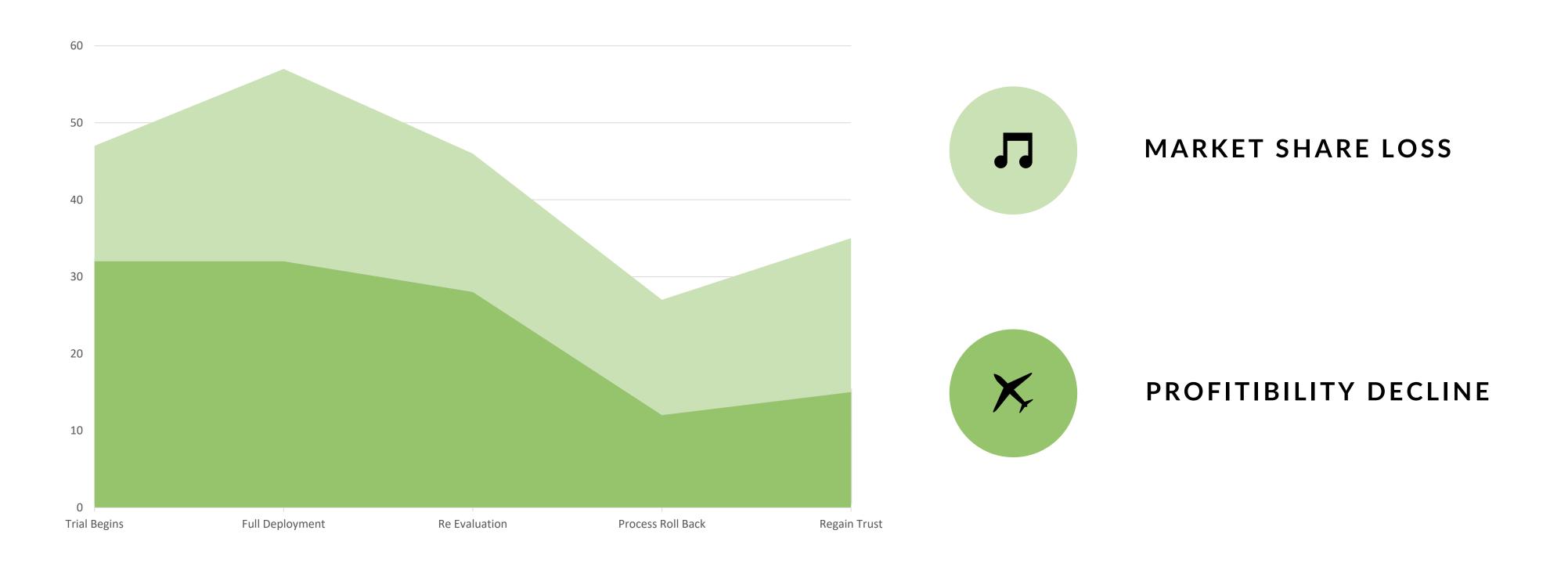
#### MACC OVERVIEW

The Midwest Architecture Community Collaboration's (MACC) purpose is to bring all domains of architecture together to share information and techniques of interest to all of us. It is our shared belief that through collaboration, we can better understand and promote the significance of architecture to business success.



### CORPORATE DISASTER

Today we will look at a company and how they used AI poorly nearly putting them out of business





# Cameron Vetter

Principal Architect

Microsoft MVP for Al

Deep Learning

Generative Al

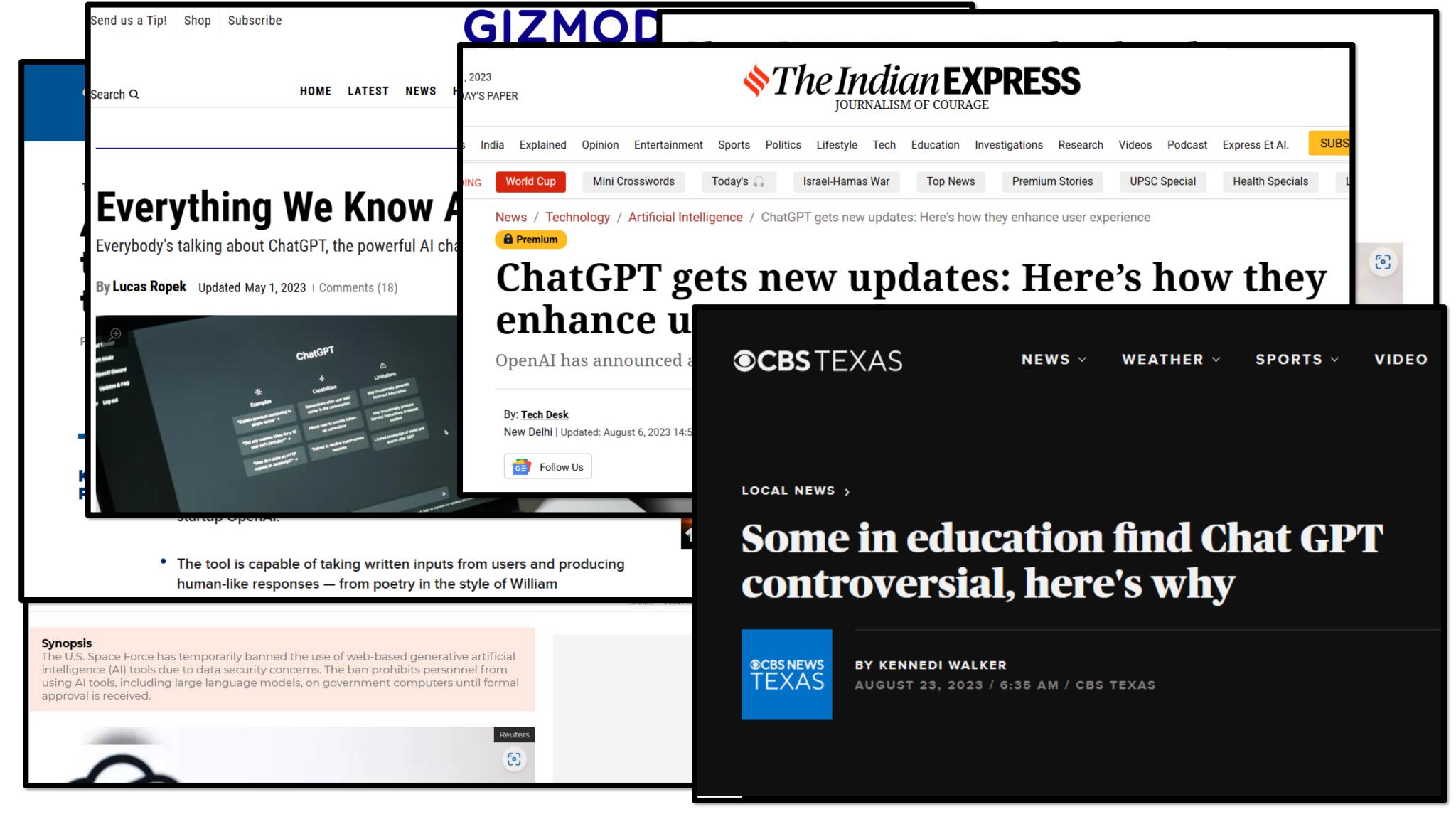
Cloud Architecture



- 1 Artificial Intelligence
- Productionalizing
  Machine Learning
- 3 Cloud Architecture
- 4 Technology Roadmaps







ChatGPT really is glorified autocomplete. It's a very cool parlor trick and fascinating from a math and linguistic perspective. But it's not sentient, and has no personality, agency, or intent. It reflects the internet's unfiltered text reflected back in a probabilistic mirror.

Given the breadth and depth of GPT-4's capabilities, we believe that it could reasonably be viewed as an early (yet still incomplete) version of an artificial general intelligence (AGI) system



THE

COMPANY:

DUNDER MIFFLIN DUNDER
MIFFLIN SELLS
PAPER



## DUNDER MIFFLIN FINALLY HAS A WEB SITE AND HAVE EMBRACED THE DIGITAL AGE





## MICHEAL SCOTT HAS DISCVORED CHAT GPT

Micheal is a the former top sales guy, but finds himself in middle management, leading the sales team. *He also has discovered the Peter Principle*.

DETAILED



# WHAT IF CHAT GPT COULD WRITE MY SALES QUOTES FOR ME?

Micheal still does sales for some of the big accounts. He loves doing sales, but hates writing the quotes.



#### **FAILURE**

It doesn't work, Micheal finds that the quotes are filled with old products and the prices aren't right.



#### **IDEA**

What about the web plugin for Chat GPT? He points that at the new web site!



#### **SUCCESS**

Suddenly the quotes use the current products and the prices are right!



#### **EMAIL SUMMARY**

I hope these ideas excite you as much as they do me, Michael! Let's show Angela and the rest of the office how we can rise above the negativity and have an amazing time together. Feel free to voice your thoughts and add any suggestions you might have.

Looking forward to your input and, most importantly, your infectious enthusiasm!

Best Regards,

Phyllis

Official Party Planner-Extraordinaire



The email from Phyllis is about planning an office party to counterbalance Angela's grumpiness. The main points are:





- 1. Acknowledging Angela's negative influence on office morale.
- 2. Proposing a "Wild West Whimsy" theme to counter Angela's strictness.
- 3. Suggesting a dancing competition to let loose and have fun.
- 4. Recommending an Office Olympics event to foster teamwork and camaraderie.
- 5. Proposing a challenge to make Angela smile, with a potential reward for the winner.
- 6. Encouraging Michael to share his thoughts and ideas for the party.

Phyllis expresses enthusiasm and hopes to create a positive and enjoyable atmosphere at the office party.

#### Dear Phyllis,

I hope this email finds you well and in high spirits. I've just read your exciting email, and I must say, I'm absolutely thrilled about your party planning enthusiasm. Your ideas sound fantastic, and I can already sense the positive energy that will come from them. We all know how Angela's rigid attitude can sometimes get the best of us, but your plan to counterbalance it with a fun-filled party is a brilliant idea!

The "Wild West Whimsy" theme is a stroke of genius. It's sure to bring out the inner cowboys and cowgirls in all of us and provide a delightful escape from Angela's strictness. I'm already picturing everyone in their cowboy hats and spurs, ready to embrace a carefree spirit.

The dancing competition idea is a hoot! I can't wait to showcase my infamous "Flailing Arms" move and see if Angela surprises us with some unexpected dance skills. Who knows, it might just be the perfect opportunity for her to let her hair down and join the fun.

The Office Olympics proposal is also a fantastic way to unite the team and show Angela that we can have a great time while maintaining our professionalism. It's a reminder that work can be enjoyable and productive simultaneously.

As for the "Smiling Challenge," it's a daring move, but it might be just the thing to crack Angela's grumpy facade. The Dundie trophy for the "Master of Angela's Smiles" is a brilliant touch, and I'm excited to see who rises to the challenge.

I'm all in for this party, Phyllis! Your infectious enthusiasm and creativity are exactly what we need to lift our spirits and create a positive atmosphere at Dunder Mifflin. Please let Angela know that I support her efforts as well, and I believe we can all come together and have a fantastic time at the party while bidding her a cheerful farewell.

Count me in for the planning, and I'll gladly share any ideas that come to mind. Let's make this party one for the books and show Angela that we know how to have fun together.

Looking forward to an amazing celebration!

#### RESPOND TO MY EMAILS

Why stop at reading my emails? Just reply to them for me as well!

#### WRITE MY **BLOG POST** FOR ME



Write a 500 word blog post about the popular paper types this fall.



Title: Embrace the Season with Popular Paper Types This Fall







As the leaves start to change colors and the air becomes crisper, fall is undeniably one of the most enchanting seasons of the year. With the change in weather comes a change in design trends, including the choice of paper types for various creative and practical projects. This fall, we're seeing a resurgence of classic paper types with a contemporary twist, bringing warmth, texture, and character to a wide range of applications.

#### 1. Kraft Paper: Rustic Elegance

Kraft paper has been a favorite choice for those who appreciate a rustic and environmentally-friendly aesthetic. This fall, it's taking center stage in many design projects. Its neutral, earthy tone complements the season's color palette, making it perfect for invitations, packaging, and even home decor. Pair it with handwritten calligraphy and warmtoned accents, and you have a combination that exudes rustic elegance.

#### 1. Cotton Paper: Luxurious Softness

Cotton paper, known for its sumptuous texture and high quality, is making waves this fall. With its soft, tactile feel, it's perfect for stationery, wedding invitations, and fine art printing. The delicate fibers and pure white appearance provide a luxurious touch, while its ability to absorb ink beautifully makes it a go-to for anyone looking to make a lasting impression.

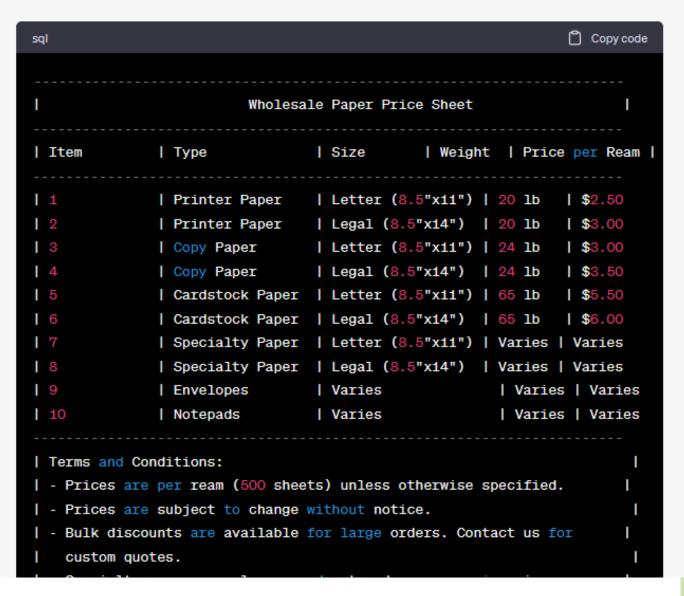
#### 1. Recycled Paper: Sustainable Choices

As environmental consciousness continues to grow, recycled paper remains a popular choice this fall. It's available in various textures and colors, making it an ideal option for ecoconscious individuals and businesses. Its versatility ensures it can be used for everything from brochures to business cards, while still being kind to the planet.

#### 1. Watercolor Paper: Artistic Flair

(5)

Creating a price sheet for wholesale paper costs will depend on various factors such as the type of paper, quantity, quality, and other variables. Below is a simplified example of a price sheet for wholesale paper costs. You can customize it to your specific needs:



## CREATE PRICING FOR NEXT QUARTER

Micheal spends SO much time figuring out the pricing for the next quarter, he just has Chat GPT do it

#### DISCOUNTS



The volume-based discounts you offer for wholesale paper purchases should be designed to incentivize larger orders while maintaining profitability. The specific discounts can vary based on your business strategy, but here are some common volume-based discount tiers that you can consider:



#### 1. Regular Price (0-9 Reams):

This is the base price that customers pay for small orders or one-time purchases.

#### 2. Small Volume Discount (10-29 Reams):

 Offer a modest discount for customers who purchase between 10 and 29 reams. For example, a 5% discount.

#### 3. Medium Volume Discount (30-49 Reams):

 Provide a more significant discount for customers ordering in medium quantities. This could be a 10% discount.

#### 4. Bulk Purchase Discount (50+ Reams):

 Encourage large orders with a substantial discount for bulk purchases. Consider a 15% discount or more for customers who buy 50 reams or more.

#### Custom Bulk Orders (100+ Reams):

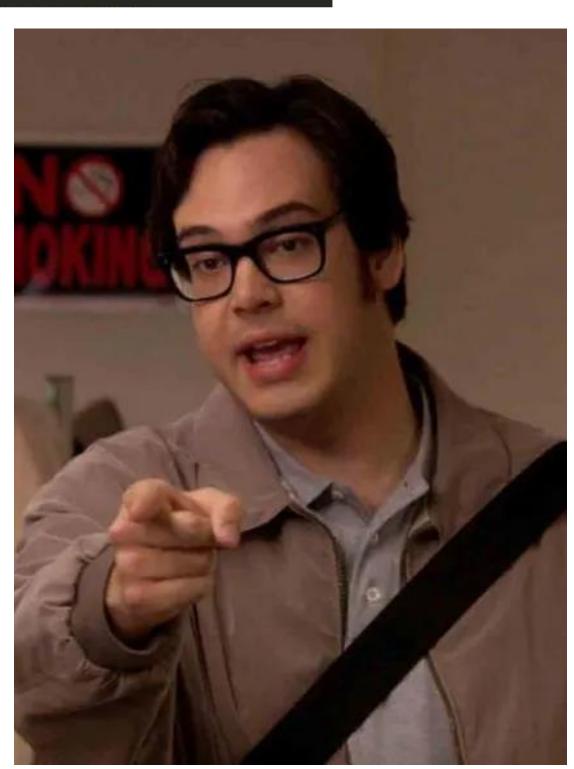
 For extremely large or recurring orders, offer custom pricing and negotiate directly with the customer. These discounts can be highly variable based on the specific needs of the customer.

#### 6. Annual/Contract Agreements:





```
le.data.users:
 client.api.statuses.user_timeline.get(scre
t', len(response.data), 'tweets from', i.scr
sponse.data) != 0:
 = response.data[0]['crea
e2 = datetime.strptime(ltd
  datetime.now()
ng = (today-ltdate2).days
wlong < daywindow:
rint i.screen_name, 'has
otaltweets += len(response
or j in response.data:
  if j.entities.urls:
       for k in j.entities
           newurl = k['exp
           urlset.add((new
rint i.screen_name, 'has
```



MEET NICK

FROM CORPORATE

> Software Team Lead

MICHEAL DEMANDS THAT NICK
INTEGRATES CHAT GPT INTO ALL
OF THEIR WORKFLOWS





It's just a bit of prompt engineering!

I can show you how if you need help...

# NICK RELENTS AND INTEGRATES CHAT GPT QUICKLY

Nick doesn't have experience with Al Models but he reads the documentation and it seems easy enough.

#### **USES THE API**

For some workflows he uses the API, the users
NEVER EVEN KNOW they are using Chat GPT.

#### MANUAL USAGE

For other more manual workflows the users get their own Chat GPT accounts and perform steps according to work instructions!

## TESTING BEGINS

The entire Scranton team comes together to help with testing...



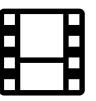
#### **TESTING GOES WELL**

Virtually every test the sales team throws at it goes well.



#### NO USER TESTING

Testing is goes so well there is no time to waste doing a limited trial with users.

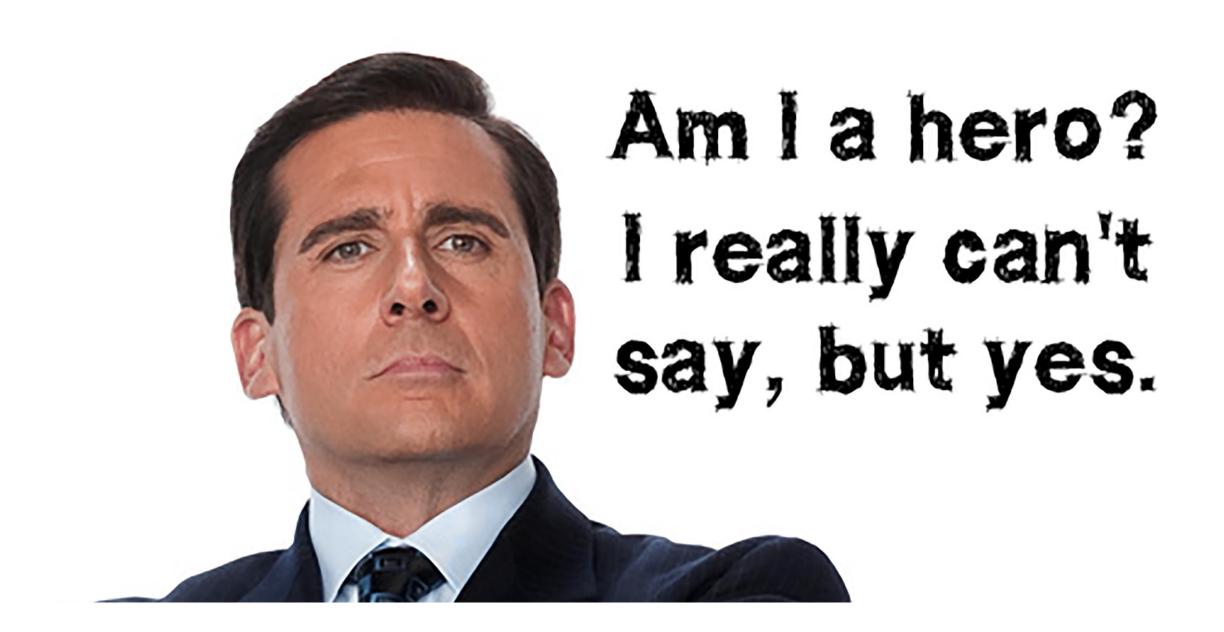


#### MINDS ARE BLOWN

The sales team starts discussing if they are going to be replaced by an Al. Dwight believes it is setient.



AGAINST NICKS ADVICE THE INTEGRATIONS GO LIVE WITH LIMITED TESTING



## THEN THE PROBLEMS START



# HALLUCINATIONS IN EMAILS

INCORRECT
INFORMATION IN
QUOTES AND
EMAILS

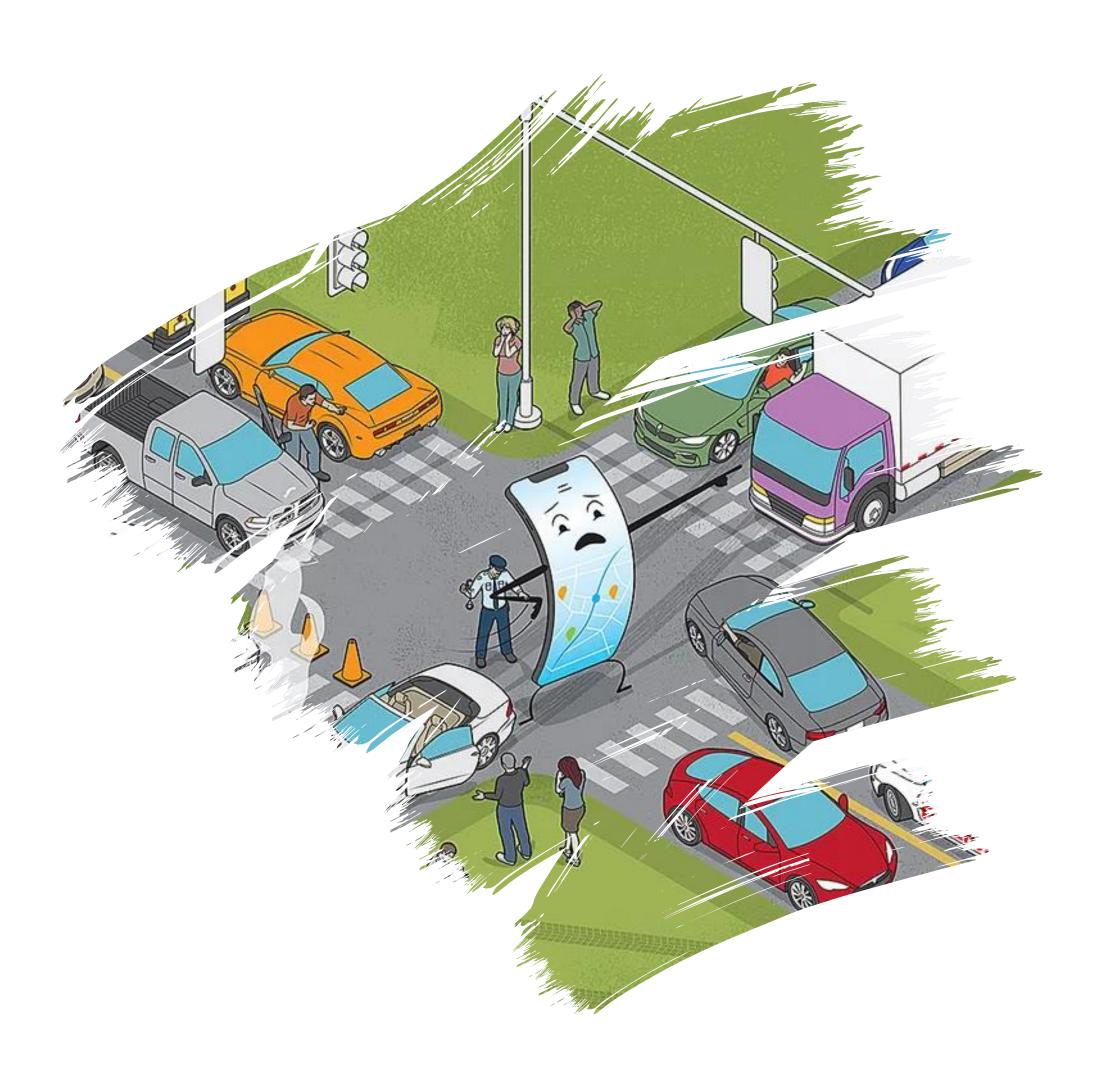




## OUTDATED INFORMATION

TOO





## CAPACITY ISSUES

# TOO EXPENSIVE

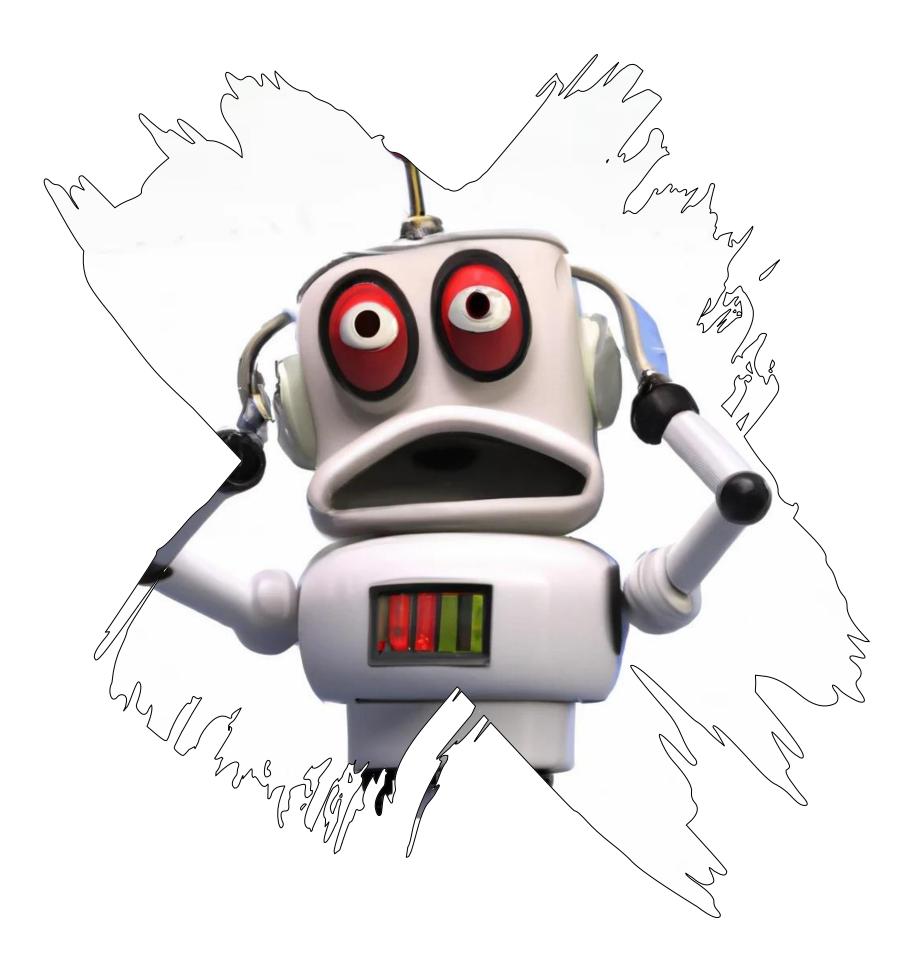




# PLAGERISM

# BIAS





# BUSINESS DOMAIN CONFUSION

CHAT GPT
CHANGED
UNEXPECTANTLY,
EVERYTHING
WORKS DIFFERENT



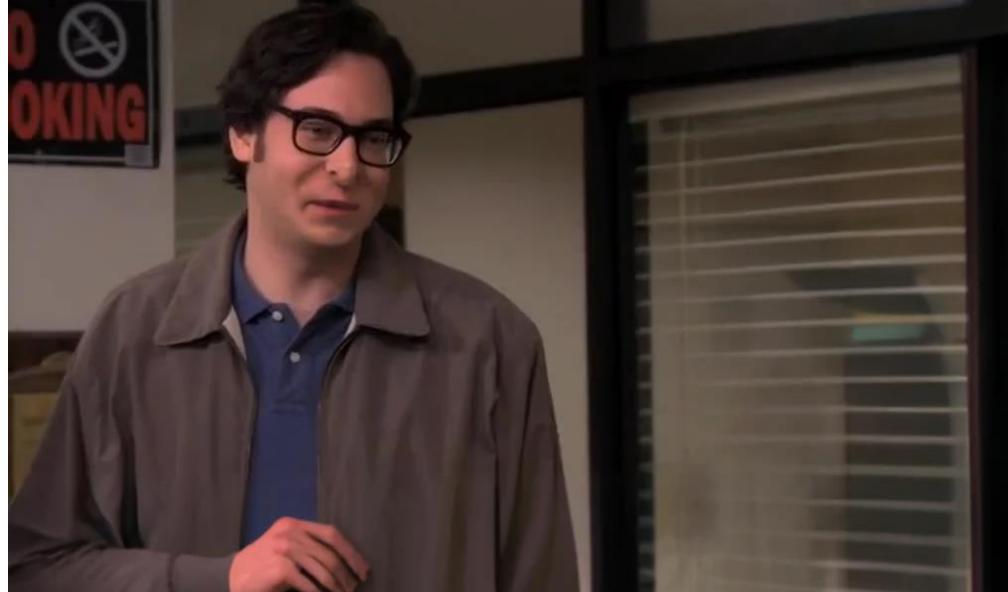
USERS
LEAKED
SECRET
COMPANY
INFO





JAILBREAK!





Our business user throws the developer under the bus and suddenly is nowhere to be found!

SCAPEGOAT



And spends the rest of the day updating his resume.

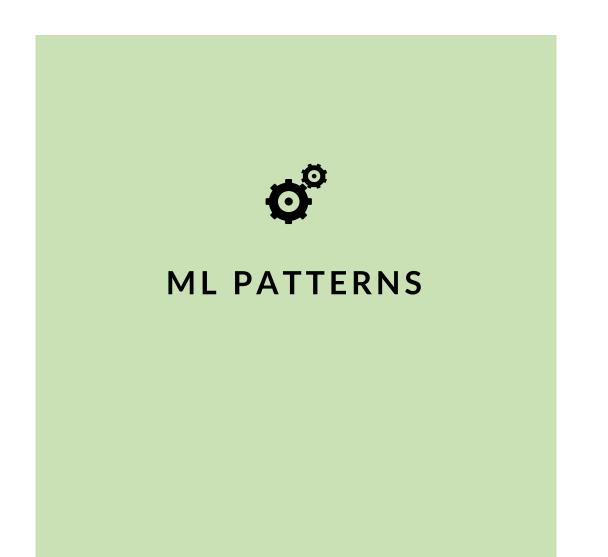
THE ARCHITECT
COMES DOWN
FROM HIS IVORY
TOWER

He proclaims *Chat GPT is a glorified tech demo* pull it from

production!









The next day the Architect proclaims, we will perform these three changes before we try again.

SOLUTION



# LLM'S ARE NOT THE SOLUTION TO EVERY PROBLEM

HAMMER



# AI ISN'T THE SOLUTION TO EVERY PROBLEM

SLEDGEHAMMER

#### ML DESIGN



#### RAG DESIGN PATTERN

Retrieval Augmented Generation.



#### FINE TUNING

Domain Adaptation through Fine Tuning.



#### PROMPT ENGINEERING

It's not just a bogus new career path.



#### MULTI MODEL SOLUTIONING

Many frameworks such as LangChain



#### MODEL SELECTION

GPT3.5, GPT 4, Bard, Llama, Flacon, PaLM, Claude, etc

# ML PATTERNS



#### **CACHING**

Completions are expensive, don't do the same ones over and over.



#### **STREAMING**

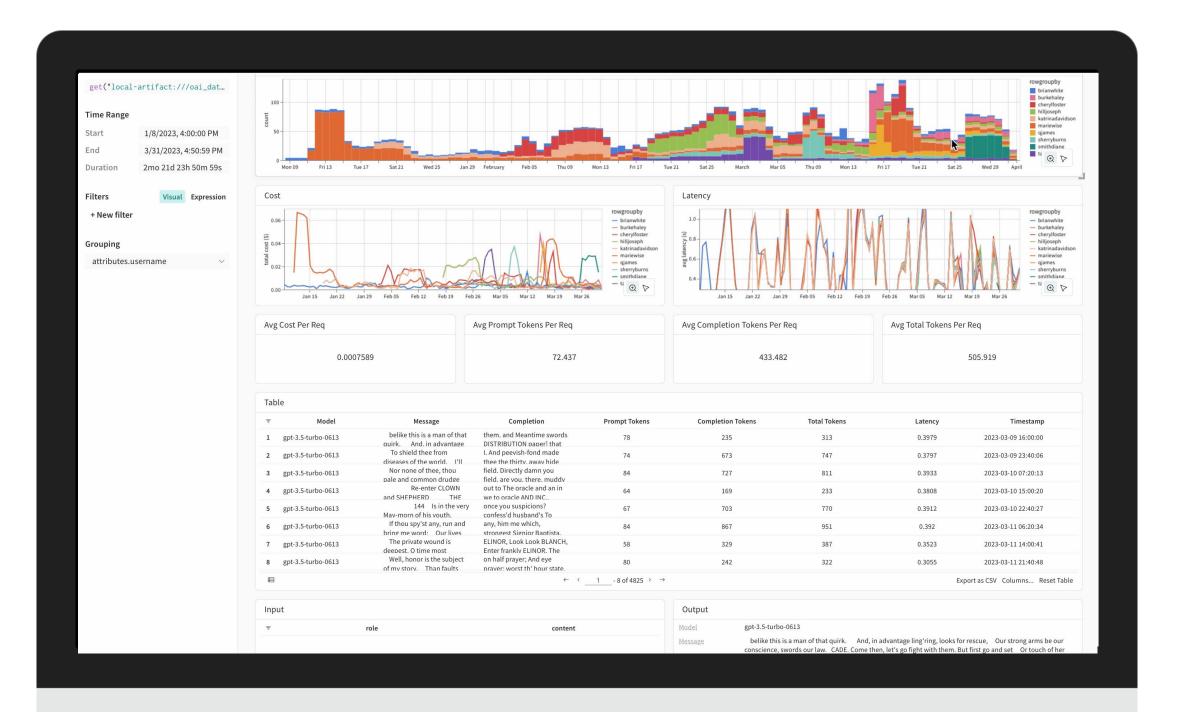
Don't make the user wait for the entire completion, stream it as it happens.



#### INSTRUMENTATION

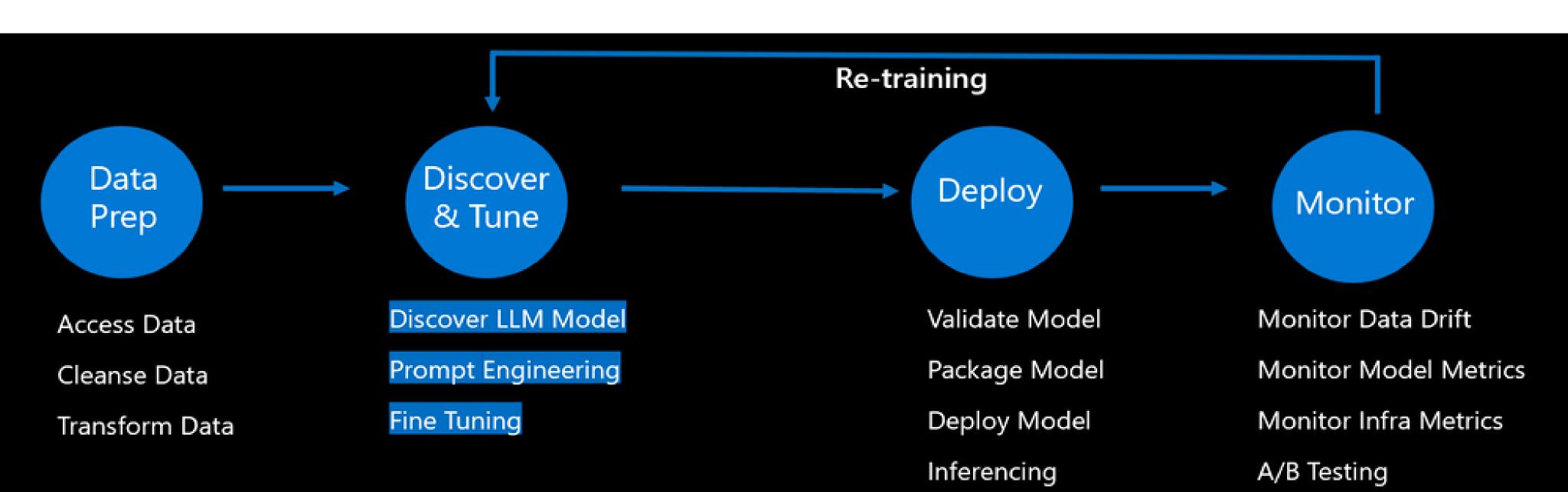
Instrument your solution and understand don't guess what is happening behind the scenes.

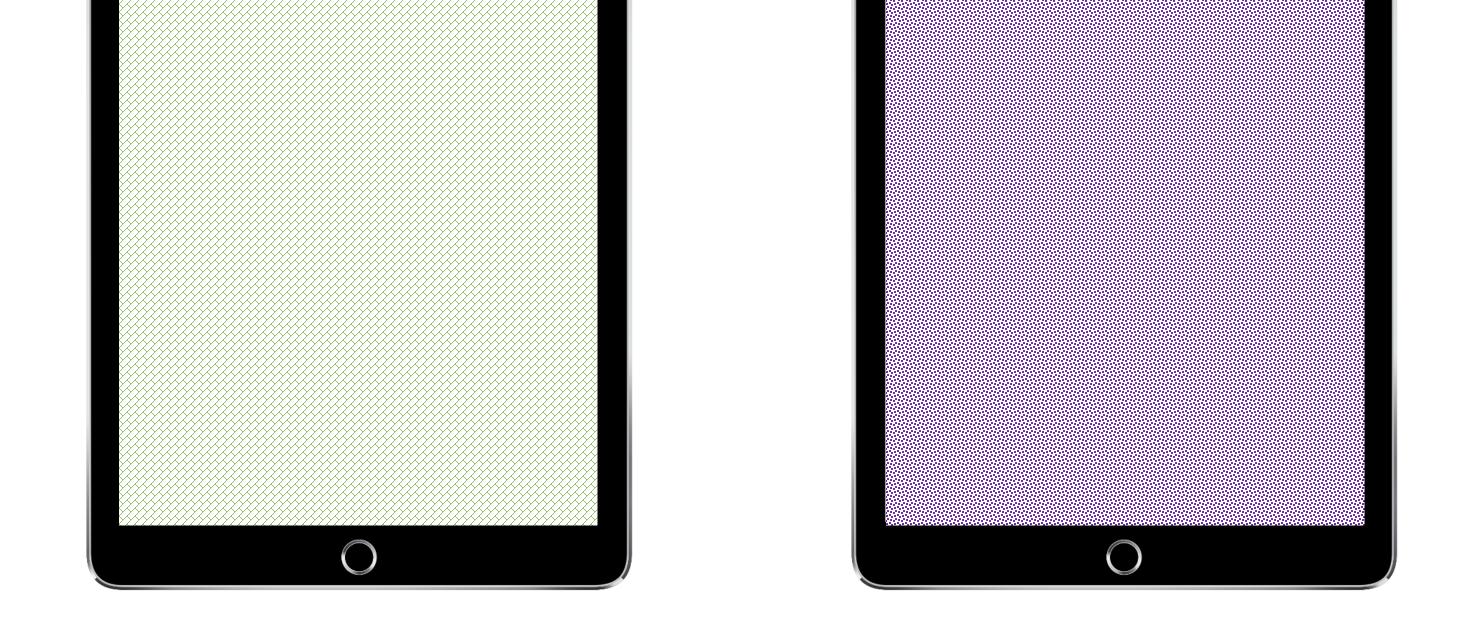
Weights &
Biases
Instrumentation



# LLM LIFECYCLE WORKFLOW

The lifecycle of a LLM integration and deployment into a product is much more complicated and requires much more care than most other software components





# A / B TESTING

Always release your model to a very small part of your audience and leave it "the old way" for the rest of your audience! Expect the unexpected when releasing any type of AI based solution!



Our hero heeds his advice and begins again.

And they all live happily ever after...

### HOW NICK CORRECTED THE ISSUES



#### **HALLUCINATIONS**

Prompt Engineering Retrieval Augmented Generation



## INCORRECT INFORMATION

Retrieval Augmented Generation
Prompt Engineering
Model Selection



# OUTDATED INFORMATION

Retrieval Augmented Generation



#### **TOO SLOW**

Model Selection

Multi Model Solutioning

Streaming

Caching



#### **CAPACITY ISSUES**

Model Selection
Streaming
Caching



## TOO EXPENSIVE

Multi Model Solutioning
Model Selection
Caching
No LLM

## HOW NICK CORRECTED THE ISSUES



#### **PLAGERISM**

Retrieval Augmented Generation Prompt Engineering



#### BIAS

Testing
Instrumentation



# BUSINESS DOMAIN CONFUSION

Fine Tuning
Retrieval Augmented Generation
Prompt Engineering



# CHAT GPT NEW RELEASE

No Chat GPT



# SECRET INFO LEAKED

Model Selection (Read the fine print)



#### **JAILBREAK**

**Prompt Engineering** 





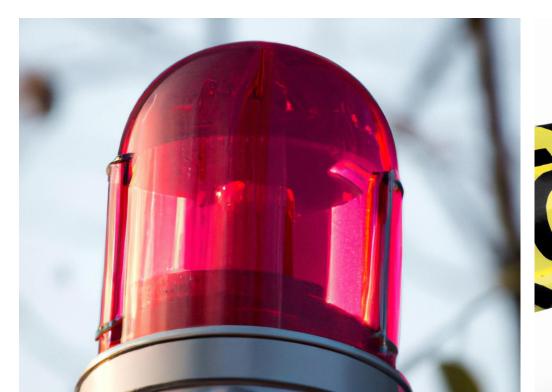




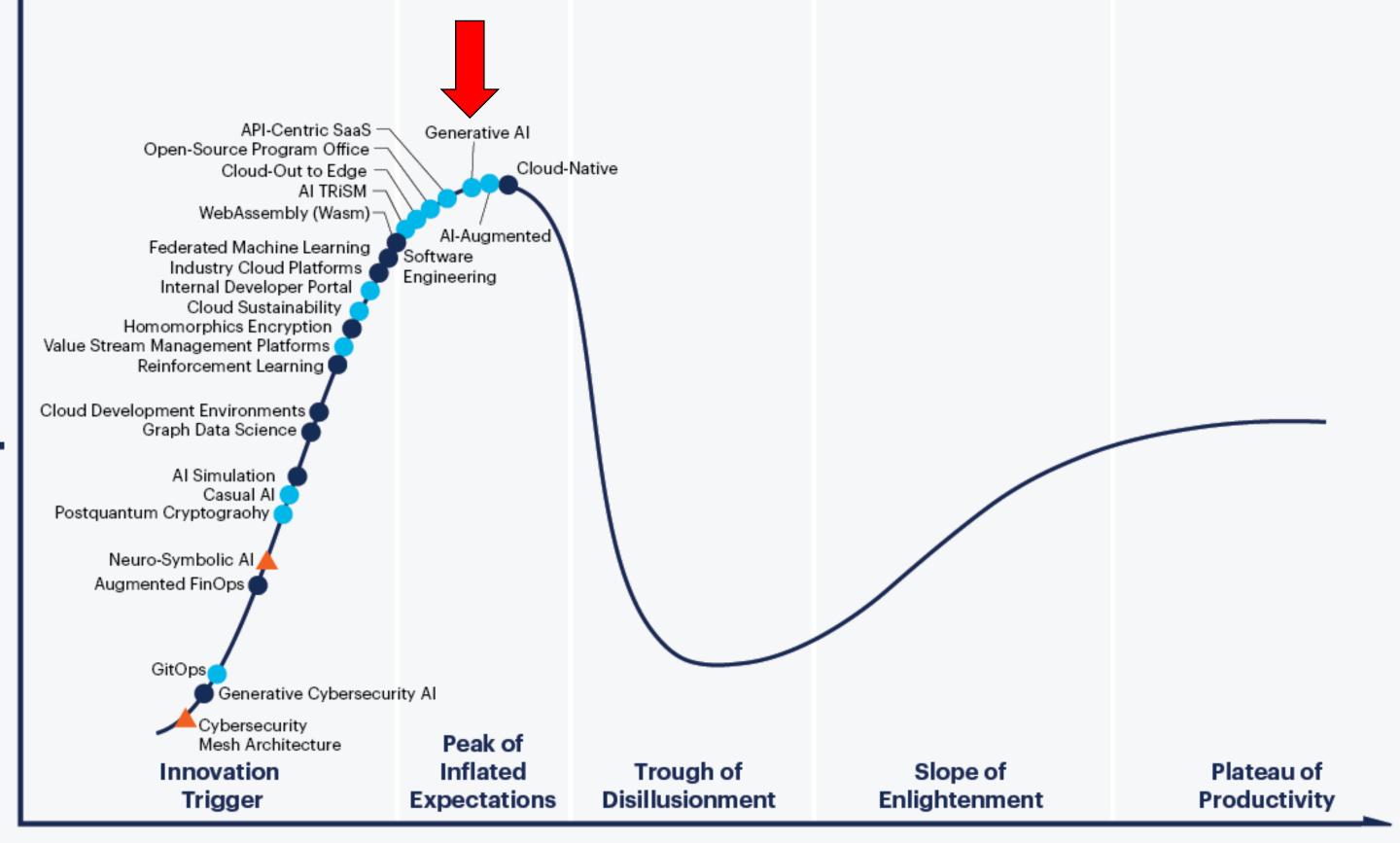
**OPENAI STATES:** "ChatGPT sometimes writes plausible-sounding but incorrect or nonsensical answers."











# Gartner Hype Cycle for Emerging Tech 2023

#### Time

Plateau will be reached:

Tread Lightly, BUT MOVE QUICKLY!

Effective use of AI will elevate your organization above the competition.

# THANK YOU!



cameron@zecilconsulting.com

www.zecilconsulting.com