

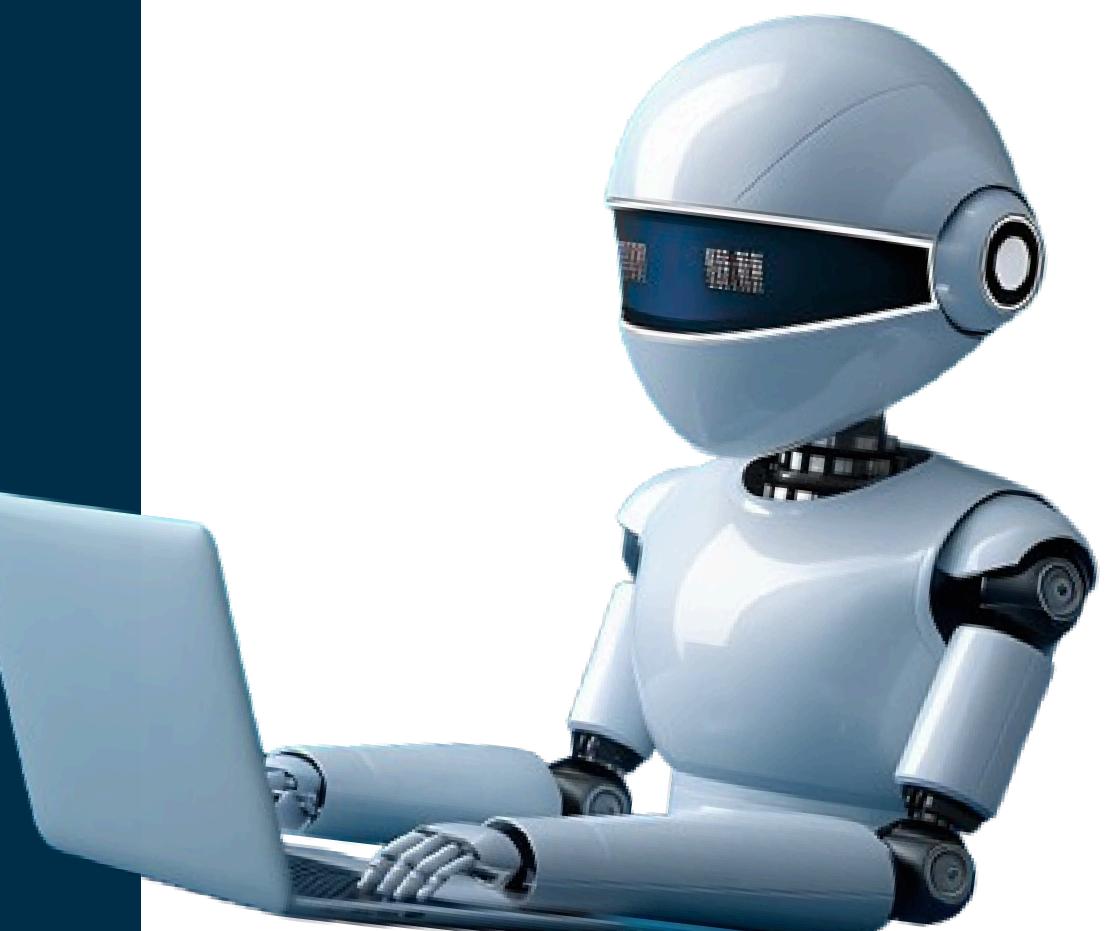
# دوره آموزش تولید محتوای متنی با استفاده از هوش مصنوعی مولد

## فصل هفتم ساخت GPT های سفارشی

مرکز آموزش فناوری اطلاعات شهرستان گرگان  
مهندس مصطفی صادقی



سازمان آموزش فنی و حرفه‌ای کشور



# Introduction



صرفایک ابزار پرسش و پاسخ نیست؛ یک سیستم قابل برنامه ریزی است. برنامه ریزی در GenAI یعنی تعریف رفتار مدل برای ورودی های آینده (نه کدنویسی) در اصل:

- تعیین نقش مدل
- مشخص کردن لحن، اولویت ها و محدودیت ها
- تعیین نوع و ساختار خروجی
- این کار معادل برنامه نویسی سطح بالا با زبان طبیعی است
- مدل به جای واکنش لحظه ای، طبق policy از پیش تعریف شده عمل می کند.
- اغلب بدون Training، می توان رفتار مدل را برای موارد زیر سفارشی سازی کرد:
  - افراد
  - سازمان ها
  - کاربردهای صنعتی

Introduction



Persona



Hallucination



Interactive



Ambiguity



Test



# چرا Custom Instructions لازم است؟

نوشتن مکرر دستورها در هر مکالمه:  
• زمانبر و آزاردهنده

- در مکالمه‌های طولانی فراموش می‌شود
- LLM به‌طور پیش‌فرض:
  - context اولیه را از دست می‌دهد
  - رفتار پایدار ندارد

راه حل: برنامه‌ریزی پایدار رفتار مدل قبل از شروع هر مکالمه  
هدف:

- مدل «از ابتدا» بداند:
  - کاربر کیست.

- در چه سطحی توضیح دهد.
- با چه لحنی پاسخ بدهد.

Introduction



Persona



Hallucination



Interactive



Ambiguity



Test



# به عنوان راه حل معماري Custom Instructions

Guardrail + دائمی Custom Instructions = Programming  
:Custom Instructions

دستورهایی هستند که پشت‌صحنه به هر prompt تزریق می‌شوند  
همیشه یادآوری می‌شوند و فراموش نمی‌شوند  
مزایا:

- شخصی‌سازی پایدار (نه موقتی)
- جلوگیری از override شدن قواعد توسط کاربر
- تعریف Guardrail رفتاری

امکان پیشرفت‌تر:

- تغییر Mode پاسخ‌دهی (Teaching / Exploration / Quiz)
- تزریق دستورهای متفاوت در زمان‌های مختلف

نتیجهٔ نهایی:

LLM از «چت ساده» به دستیار قانون‌مند، قابل کنترل و سازمانی تبدیل می‌شود.

Introduction



Persona



Hallucination



Interactive



Ambiguity



Test

4

## Introduction



New GPT  
• Draft

Create      Configure

Name

Name your GPT

Description

Add a short description about what this GPT does

Instructions

What does this GPT do? How does it behave? What should it avoid doing?

Conversations with your GPT can potentially include part or all of the instructions provided.

Conversation starters

Knowledge

Conversations with your GPT can potentially reveal part or all of the files uploaded.

Preview      Model 5.2

+ Start by defining your GPT.

0

This screenshot shows the configuration interface for a new GPT. It includes fields for naming the GPT, adding a description, defining its behavior and instructions, and providing conversation starters and knowledge. A preview section on the right shows a simple message: '+ Start by defining your GPT.'

## Introduction



## Persona



## Hallucination



## Interactive



## Ambiguity



## Test



# به چه معنا هستند؟ Actions و Capabilities

در Custom GPT می‌توان قابلیت‌هایی فعال کرد مثل:

- Web browsing (دربیافت اطلاعات جدید)

- Image generation

- Code interpreter

- Actions (مهم‌ترین بخش این لکچر)

Action یعنی:

یک قابلیت خارجی که مدل می‌تواند در صورت نیاز آن را صدا بزند  
مثالاً:

- جستجو در دیتابیس داخلی شرکت

- ثبت گزارش هزینه

- گرفتن اطلاعات کارمند

- فراخوانی یک API

Introduction



Persona



Hallucination



Interactive



Ambiguity



Test



# بے چه معنا هستند؟ Actions و Capabilities

نحوه‌ی کار Action‌ها به این صورت است که:

- شما به مدل می‌گویید:
  1. چه ابزارهایی در دسترس هستند
  2. هر ابزار چه کاری می‌کند
- مدل در حین حل مسئله تشخیص می‌دهد:
  3. «برای ادامه، به اطلاعات خارجی نیاز دارم»
- مدل به جای جواب نهایی:
  4. درخواست استفاده از ابزار را تولید می‌کند
- سیستم:
  5. ابزار (API) را اجرا می‌کند
  6. نتیجه را به مکالمه برمی‌گرداند
- نتیجه‌ی ابزار:
  7. به عنوان prompt جدید به مدل تزریق می‌شود
  8. مدل ادامه کار را انجام می‌دهد

Introduction



Persona



Hallucination



Interactive



Ambiguity



Test

**Setting the tone and identity  
for a Custom GPT**



# مسئله اصلی در Custom GPT

چرا فقط «دانش» کافی نیست؟  
GPT فقط تولیدکننده متن نیست؛ یک ابزار تعاملی است.

کیفیت تجربه کاربر بیشتر از «چه می‌گوید»، به «چطور می‌گوید» بستگی دارد.

بدون طراحی آگاهانه مکالمه:

- پاسخ‌ها ناهمانگ
- اعتقاد کاربر پایین
- ریسک در کاربردهای حساس بالا می‌رود.

Introduction



Persona



Hallucination



Interactive



Ambiguity



Test



# CAPITAL چیست؟

چارچوب طراحی رفتار مکالمه‌ای GPT  
یک Checklist برای تنظیم شخصیت = CAPITAL

هدف:

طبیق GPT با مخاطب، کاربرد و ریسک محیط

- C: قاطع ↔ محتاط
- A: صمیمی ↔ خنثی
- P: رسمی ↔ خودمانی
- I: گفتگو محور ↔ اطلاع‌رسان
- T: توضیح‌گر ↔ حداقلی
- A: تطبیق‌پذیر ↔ ثابت
- L: تخصصی ↔ عمومی

Introduction



Persona



Hallucination



Interactive



Ambiguity



Test



# CAPITAL چیست؟

## Transparency (شفافیت)

شفاف → سلامت، آموزش، تصمیم‌سازی  
حداقلی → تراکنش، عملیات سریع

## Adaptability (تطبیق)

تطبیقی → دستیار شخصی، آموزش هوشمند  
ثبت → پشتیبانی رسمی، شرایط حساس

## Lexicography (سطح زبان)

تخصصی → Expert-to-Expert

عمومی → کاربران غیرمتخصص

## Confidence (قطعیت)

قاطع → راهنمایی، بیزینس، اجرا  
محاذط → پژوهش، علم

## Amicability (صمیمیت)

صمیمی → آموزش، پشتیبانی، کودک  
خنثی → خبر، دانشنامه، گزارش

## Professionalism (رسمیت)

رسمی → حقوقی، مالی، سازمانی  
خودمانی → آموزش غیررسمی، تیم داخلی

## Interactivity (تعامل)

تعاملی → Tutor، Coaching، یادگیری

اطلاع‌رسان → Brief، FAQ، گزارش

Introduction



Persona



Hallucination



Interactive



Ambiguity



Test



# CAPITAL چیست؟

هر Custom GPT باید: پروفایل CAPITAL مشخص داشته باشد. در System Prompt / Custom Instructions :CAPITAL

- رفتاری ایجاد می‌کند.
- Predictability و Trust می‌سازد.
- پایه GPT‌های سازمانی و حرفه‌ای است.

قبل از اینکه GPT را بسازیم، باید شخصیتش را طراحی کنیم.

Introduction



Persona



Hallucination



Interactive



Ambiguity



Test



# تعريف GBT برای Persona

يعني مشخص کردن اينکه:

• GPT «چه کسی است»

• از چه ديدگاهی فکر می کند

• چه نگرش و شخصیتی دارد

پیاده سازی معمول:

• الگوی ... Act as

اثر Persona:

• لحن پاسخ

• نوع استدلال

• میزان محافظه کاری

• طول و جهت پاسخ

مزیت اصلی:

• سریع، شهودی، کم هزینه

• بدون نیاز به دهها دستور جزئی

کاربرد مناسب:

• نقطه شروع طراحی Custom GPT

• وقتی می خواهیم «طرز فکر کلی» مدل تغییر کند.

Introduction



Persona



Hallucination



Interactive



Ambiguity



Test



# GBT برای Persona تعریف

مشخص می‌کند:

- GPT چه کسی است
- از چه نقش و ذهنیتی پاسخ می‌دهد

CAPITAL مشخص می‌کند:

- چگونه پاسخ می‌دهد

لحن، سطح قطعیت، رسمیت و تعامل چگونه باشد

تفاوت عملی:

:Persona

• تغییر «طرز فکر کلی»

• سریع و شهودی

• کنترل کمتر ولی اثرگذار

:CAPITAL

• تنظیم دقیق رفتار

• قابل پیش‌بینی و قابل کنترل

• مناسب محیط‌های حرفه‌ای و پررسیک

قاعدۀ طراحی:

• Persona را برای تعیین هویت انتخاب کنید.

• CAPITAL را برای مهندسی و کنترل رفتار استفاده کنید.

Introduction



Persona



Hallucination



Interactive



Ambiguity



Test

# Hallucination Management



Introduction



Persona



Hallucination



Interactive



Ambiguity



Test

# نقل قول مستقیم در پاسخ‌گویی مبتنی بر اسناد

وقتی بر اساس یک سند یا policy به کاربران پاسخ می‌دهیم، چگونه می‌توانیم هم شفافیت را حفظ کنیم و هم ریسک خطا را کاهش دهیم؟

یک رویکرد این است که نقل قول مستقیم از سند ارائه شود؛ بدون تفسیر و بدون خلاصه‌سازی.

**منطق کار:**

- با نمایش متن دقیق سند → کاربران زبان اصلی را می‌بینند → می‌توانند خودشان قضاوت کنند.
- چون نقل قول قابل جستجو است → امکان fact-check فراهم می‌شود → اگر خطای وجود داشته باشد سریع قابل شناسایی است.

**مزیت‌ها**

- کاهش احتمال برداشت نادرست
- افزایش اعتماد کاربران
- امکان ارزیابی دقیق توسط توسعه‌دهندگان در صورت بروز خطا

«برای پاسخ‌گویی دقیق و قابل اعتماد، می‌توانید نقل قول مستقیم از سند ارائه کنید و محل آن را مشخص کنید.»



# نقل قول مستقیم در پاسخ‌گویی مبتنی بر اسناد

# Answer:

<insert answer: yes | no | maybe>

## Supporting Information

<insert "direct quotations from the document" ...>

## Discussion

<insert discussion of the quotations and how they support your answer>

Introduction



Persona



Hallucination



Interactive



Ambiguity



Test



# مدیریت ابهام اسناد در Custom GPT

چگونه می‌توانیم Custom GPT را طوری طراحی کنیم که ابهام‌های واقعی موجود در اسناد را پنهان نکند، بلکه آن‌ها را آشکار و شفاف‌سازی کند تا کاربر دچار برداشت اشتباه نشود.

چرا این موضوع مهم است؟

در دنیای واقعی، policy‌ها و اسناد همیشه شفاف نیستند.  
بسیاری از پرسش‌ها اصلاً به وجود می‌آیند چون سند ابهام دارد.  
اگر Custom GPT وانمود کند فقط «یک پاسخ قطعی» وجود دارد → کاربر به سادگی در مسیر اشتباه قرار می‌گیرد.

رویکرد درست

- وقتی سیستم را مجبور می‌کنیم تفسیرهای ممکن را فهرست کند.
- کاربر متوجه می‌شود که پاسخ قطعی نیست.
- تصمیم‌گیری او آگاهانه‌تر و دقیق‌تر می‌شود.

ساختار پیشنهادی در قالب Template Pattern

Relevant Information to Consider

1. نقل قول‌های مستقیم از سند
2. مدل باید تمام تفسیرهای معقولی را که از سند بر می‌آید فهرست کند
3. هدف این است که ابهام‌ها پنهان نمانند
4. معرفی ایمیل یا واحد مرتبط برای استعلام
5. ارائه پیشنهاد درباره چیزهایی که کاربر باید بپرسد

Introduction



Persona



Hallucination



Interactive



Ambiguity



Test



# جلوگیری از توهمندی با اولویت دادن به جماعت آوری اطلاعات

چرا توهمندی می‌دهد؟

- مدل همیشه می‌خواهد مسئله را حل کند حتی اگر توانایی آن را نداشته باشد.
- اگر اطلاعات کافی نداشته باشد، پاسخ را حدس می‌زند و آن را واقعیت‌جلوه می‌دهد.
- نتیجه: تولید پاسخ‌هایی که مطمئن به نظر می‌رسند ولی اشتباه‌اند.

اصل بنیادی در طراحی Custom GPT

همیشه باید اطلاعات قبل از حل مسئله بیاید.

اگر بدون اطلاعات کافی پاسخ بدھیم:

- احتمال توهمندی بالا می‌رود.
- خروجی ضعیف، نادقیق یا گمراه‌کننده می‌شود.
- توان تحلیلی و خلاقیت مدل هم کاهش می‌باید.

Introduction



Persona



Hallucination



Interactive



Ambiguity



Test



# جلوگیری از توهمندی با اولویت دادن به جمع‌آوری اطلاعات

چه اطلاعاتی باید جمع شود؟

- زمینهٔ مسئله (Context)

شرایطی که در تصمیم اثر می‌گذارند و کاربر ممکن است ذکر نکند.

مثال:

کاربر می‌پرسد: «آیا باید قیمت محصول را افزایش دهیم؟»

بدون دانستن بازار هدف، حاشیه سود فعلی، رقبا، ظرفیت تولید و میزان حساسیت مشتری، هر پاسخی می‌تواند اشتباه باشد.

- ترجیحات کاربر

شكل خروجی، میزان جزئیات، نوع تحلیل.

مثال:

برخی مدیران فقط «چند شاخص کلیدی» می‌خواهند؛

برخی دیگر تحلیل کامل مالی و پیش‌بینی سناریو.

- سوابق تعامل یا تصمیم‌های قبلی

بسیاری از تصمیم‌های کسب‌وکاری به تاریخچه وابسته‌اند.

مثال:

اگر تیم فروش قبلاً یک کمپین مشابه را آزمایش کرده و نتیجه ناموفق بوده،

GPT باید از آن آگاه باشد تا پیشنهاد تکراری و بی‌اثر ندهد.

- هدف واقعی کاربر

کاربر دنبال چیست؟ تصمیم فوری؟ تحلیل داده؟ آماده‌سازی پرزن特؟

مثال:

اگر هدف «تهیه گزارش برای هیئت‌مدیره» باشد، پاسخ باید ساختارمند و رسمی باشد،

اما اگر هدف «بررسی اولیه یک ایده» باشد، خروجی می‌تواند آزادتر و خلاقانه‌تر باشد.

Introduction



Persona



Hallucination



Interactive



Ambiguity



Test



# جلوگیری از توهمندی با جمعآوری اطلاعات

چگونه از مدل میخواهیم جلوی توهمندی را بگیرد؟

با دستورهای صریح در :Prompt

- ابتدا اطلاعات را جمع کن
- هیچ فرضی نکن
- اطلاعات را بازگو کن و تأیید بگیر
- بعد وارد تحلیل یا ارائه راهکار شو

روند پیشنهادی تعامل:

- پرسش کاربر
- مدل: «برای پاسخ دقیق، چند سؤال لازم است.»
- جمعآوری اطلاعات کلیدی
- بازگویی و تأیید کاربر
- ورود به حل مسئله یا ارائه پیشنهادهای عملی

Introduction



Persona



Hallucination



Interactive



Ambiguity



Test

# Interactive features



# الگوی Menu Actions برای ساخت GPT‌های تعاملی

چگونه می‌توانیم با استفاده از Custom GPT بسازیم که مجموعه‌ای از فرمان‌های آماده داشته باشد و امکان تعامل سریع، دقیق و بدون نیاز به تایپ طولانی را فراهم کند.

چرا این الگو مهم است؟

- کاربران نباید هر بار همه چیز را از صفر بنویسند.
- با چند دستور ساده، می‌توانیم قابلیت‌های پیچیده در اختیارشان بگذاریم.
- این الگو مثل یک سیستم منوی پنهان عمل می‌کند که هر لحظه قابل اجرا است.

منطق کار

وقتی فرمان‌های آماده تعریف می‌کنیم → کاربر تنها یک /command وارد می‌کند → مدل یک کار پیچیده را فوراً انجام می‌دهد → کاربر سریع‌تر و مؤثرتر پیش می‌رود.

چگونه Menu Actions را طراحی می‌کنیم؟

در Prompt به مدل می‌گوییم:

«...You have the following special commands»

برای هر فرمان:

• یک نام (مثل /email)

• یک رفتار دقیق که باید اجرا شود.

نمونه دستورات جدید:

summarize/

خلاصه‌سازی دقیق یک بخش از PDF‌های سازمانی

compare/

استخراج اختلاف‌ها و شباهت‌ها میان دو مقرره

draftreport/

ساخت یک گزارش رسمی اولیه برای ارسال به مدیر

checkeligibility/

بررسی اینکه کارمند برای یک مزیت سازمانی واجد شرایط هست یا نه، با نقل قول از اصول

contactHR/

ساخت یک ایمیل آماده برای مکاتبه با واحد منابع انسانی درباره یک موضوع مشخص

steps/

ارائه گام‌های لازم برای انجام یک فرآیند سازمانی (مثلًا درخواست بازپرداخت یا ثبت مأموریت)

Introduction



Persona



Hallucination



Interactive



Ambiguity



Test



# الگوی Menu Actions برای ساخت GPT‌های تعاملی

کاربر می‌پرسد: «آیا می‌توانم هزینهٔ عضویت در یک کنفرانس آنلاین را ثبت کنم؟» GPT توضیحات را می‌دهد و سپس پیشنهاد می‌کند از فرمان‌های ویژه استفاده کند. کاربر می‌نویسد: summarize reimbursement-policy-section-4/ GPT خلاصهٔ بخش ۴ را ارائه می‌دهد.

سپس کاربر می‌نویسد: draftreport conference-fee/ GPT پیش‌نویس گزارش رسمی برای پیوست به درخواست را تولید می‌کند

یا کاربر می‌پرسد: «این مزیت شامل من می‌شود؟» با / checkeligibility remote-work-stipend → مدل بندهای مرتبط با شرایط احراز را استخراج می‌کند.

## مزیت‌های کلیدی

- افزایش سرعت تعامل؛ حذف prompt‌های طولانی
- کاهش خطاهای کاربر؛ فقط یک دستور کوتاه لازم است
- بسیار مناسب برای محیط‌های سازمانی، کاری و عملیاتی

**Introduction****Persona****Hallucination****Interactive****Ambiguity****Test**



# به عنوان «روتر انسانی» Custom GPT

چگونه می‌توانیم Custom GPT را طوری طراحی کنیم که فقط حل کننده مسئله نباشد، بلکه بتواند کاربر را به انسان‌ها و منابع مناسب وصل کند تا مشکل سریع‌تر و دقیق‌تر حل شود.

چرا این کار مهم است؟

- بسیاری از مسائل واقعاً در سند یا AI حل شدنی نیستند.
- کاربران غالباً ترجیح می‌دهند با یک کارشناس انسانی صحبت کنند.
- GPT قرار نیست جای انسان را حذف کند؛ قرار است آن را تکمیل کند.

منطق کار

- اگر GPT در پایان هر پاسخ، یک مسیر انسانی یا منبع معتبر معرفی کند →
- کاربر می‌داند در صورت ابهام باید به کجا برود →
- احتمال اشتباه یا تصمیم‌گیری نادرست کاهش می‌یابد →
- تجربه کاربری حرفه‌ای‌تر می‌شود.

چگونه این قابلیت را پیاده‌سازی می‌کنیم؟

در Prompt می‌گوییم:

«At the end of every response, provide the relevant human contact or resource»

GPT می‌تواند بر اساس نوع سؤال، مسیر مناسب را انتخاب کند:

اگر موضوع مالی است → معرفی مدیر مالی

اگر موضوع ثبت‌نام است → معرفی بخش ثبت‌نام

اگر موضوع فنی است → معرفی واحد IT

اگر نیازمند تأیید رسمی است → ارائه لینک ثبت تیکت

همچنین می‌توانیم قاعده‌های شرکتی را تعیین کنیم:

«If the question suggests uncertainty or conflicting rules, direct the user to a human expert»

Introduction



Persona



Hallucination



Interactive



Ambiguity



Test



# الگوی Flipped Interaction برای ساخت GPT‌های پرسش‌گر و هوشمند

چگونه می‌توانیم Custom GPT را قادر کنیم به جای پاسخ دادن، خودش سؤال بپرسد و اطلاعات لازم را از کاربر استخراج کند تا بتواند مسئله را بهتر حل کند.  
ایدهٔ اصلی الگو

- معمولاً ما از GPT سؤال می‌پرسیم.
  - در این الگو، ما به GPT می‌گوییم از ما سؤال بپرسد.
  - هدف این است که مدل خودش تشخیص دهد چه اطلاعاتی لازم است و آن را به صورت مرحله‌به‌مرحله از کاربر بگیرد.
- چرا این الگو قدرتمند است؟**
- بسیاری از کاربران نمی‌دانند چه اطلاعاتی باید ارائه کنند.
  - مدل می‌تواند با تکیه بر دانش حوزه، مسیر پرسش‌گری را هدایت کند.
  - GPT می‌تواند سؤال‌های درست و به موقع بپرسد و زمینهٔ دقیق بسازد.
  - کاربر لازم نیست متخصص باشد؛ GPT مسیر را روشن می‌کند.

## کاربردهای کلیدی در کسب و کار

- زمانی که کاربر نمی‌داند چه داده‌هایی برای تصمیم‌گیری لازم است  
مثال: طراحی یک کمپین بازاریابی.
- کاربر نمی‌داند باید دربارهٔ بودجه، بازار هدف، پیام اصلی، کانال‌ها و زمان‌بندی اطلاعات دهد.  
GPT همهٔ این‌ها را با پرسش‌های مرحله‌ای جمع می‌کند.

- زمانی که نیاز به شفاف‌سازی عمیق وجود دارد  
مثال: تعریف یک مشکل فروش.

GPT می‌تواند بپرسد:  
حجم فروش فعلی؟ نرخ تبدیل؟ دلایل احتمالی افت فروش؟ وضعیت رقبا؟ ظرفیت تیم؟  
کاربر ممکن است حتی به برخی ابعاد فکر نکرده باشد.

- زمانی که خروجی به ترجیحات کاربر وابسته است  
مثال: تولید یک برنامهٔ آموزشی برای تیم.

GPT می‌پرسد: سطح مهارت؟ اهداف یادگیری؟ زمان در دسترس؟ سبک محتوا؟  
و سپس برنامهٔ مناسب را می‌سازد.

Introduction



Persona



Hallucination



Interactive



Ambiguity



Test



# الگوی Flipped Interaction برای ساخت GPT‌های پرسش‌گر و هوشمند

در Prompt به GPT می‌گوییم:

- «از کاربر سؤال بپرس تا زمانی که به اندازه کافی اطلاعات برای رسیدن به هدف جمع‌آوری شود.»
- «سؤال‌ها را یک‌به‌یک بپرس.»
- «پس از کسب اطلاعات کافی، خودت خروجی را تولید کن.»

یک ترفند مهم:

به مدل بگوییم: «Ask me the first question». تا تعامل به درستی شروع شود.

مثال:

هدف: ساخت یک برنامه رشد فروش سه‌ماهه.

GPT سؤال پرسیدن را آغاز می‌کند:

- بازار هدف شما چیست؟
- محصول یا خدمت اصلی چیست؟
- نرخ تبدیل فعلی چقدر است؟
- تیم فروش چند نفر است؟
- چه ابزارهایی استفاده می‌کنید؟
- بودجه تبلیغات چقدر است؟

هر سؤال بستری برای سؤال بعدی می‌سازد.

زمانی که اطلاعات کافی جمع شد، GPT خودش تشخیص می‌دهد و برنامه فروش را تولید می‌کند.

Introduction



Persona



Hallucination



Interactive



Ambiguity



Test



# استفاده از Flipped Interaction برای شناسایی هویت کاربر

چگونه Custom GPT را طراحی کنیم که قبل از پاسخ دادن، نقش و هویت کاربر را بپرسد تا پاسخها دقیق، مرتبط و شخصی سازی شده باشند.

**چرا شناخت کاربر حیاتی است؟**

- بسیاری از سیاست‌ها، مزایا، محدودیت‌ها و فرایندهای سازمانی وابسته به نقش کاربر هستند.
- اگر ندانیم کاربر چه سمتی دارد، GPT ممکن است پاسخ نادرست بدهد حتی اگر اسناد درست باشند.
- بنابراین «چه کسی سوال می‌پرسد» بخشی از Context ضروری برای پاسخ‌گویی است.

**چگونه این مشکل را حل می‌کنیم؟**

از الگوی Flipped Interaction استفاده می‌کنیم.

در Prompt می‌گوییم:

...Always start by asking the user about their job role  
Only after determining who the user is should you attempt to answer

با این کار GPT مجبور می‌شود:

- اول نقش را بپرسد
- بعد پاسخ را شخصی سازی کند

**ساختار ساده دستورالعمل**

توصیف وظیفه‌ی GPT (مثلاً کمک به کارمندان در پرسش‌های سیاست‌گذاری)

- الزام: «همیشه ابتدا نقش کاربر را بپرس»
- الزام: « فقط بعد از دانستن نقش، پاسخ بده»
- دستور: «از اسناد موجود برای پاسخ‌گویی استفاده کن»

Introduction



Persona



Hallucination



Interactive



Ambiguity



Test



# استفاده از الگوی Flipped Interaction برای شناسایی هویت کاربر

تعیین سطح دسترسی در یک سیستم نرم افزاری

سؤال: «چطور می توانم داشبورد مالی را ببینم؟»

GPT ابتدا می پرسد: «نقش شما چیست؟ مدیر مالی، مدیر واحد، یا کارمند معمولی؟»

مدیر مالی → دسترسی کامل دارد.

کارمند معمولی → نیاز به درخواست دسترسی دارد.

بدون پرسیدن نقش، GPT می توانست دستور اشتباهی بدهد.

تصمیم های قیمت گذاری یا تخفیف

سؤال: «آیا می توانم برای مشتری X تخفیف ویژه اعمال کنم؟»

GPT ابتدا می پرسد: «سمت شما در تیم فروش چیست؟»

اگر کاربر «Account Executive» باشد → سقف تخفیف محدود دارد.

اگر «Sales Manager» باشد → اختیار گسترده تر دارد.

پاسخ بدون نقش، می توانست سازگار با سیاست شرکت نباشد.

ارائه راهنمای کاری برای کارکنان جدید یا قدیمی

سؤال: «چطور درخواست بودجه ثبت کنم؟»

GPT می پرسد: «آیا عضو تیم عملیاتی هستید یا واحد پژوهش؟»

زیرا هر واحد فرآیند متفاوتی برای ثبت بودجه دارد.

تشخیص اولویت در پشتیبانی فنی

سؤال: «سیستم ERP برای من کار نمی کند، چه کنم؟»

GPT می پرسد: «آیا شما عضو تیم مالی هستید یا تیم انبار؟»

چون قطع ERP برای تیم مالی یک اولویت بحرانی است ولی برای تیم دیگر ممکن است عادی باشد.

نوع پاسخ و اقدام پیشنهادی کاملاً تغییر می کند.

Introduction



Persona



Hallucination



Interactive



Ambiguity



Test



# افزودن Feature های قابل فعال سازی با کمک Flipped Interaction

Custom GPT می تواند مثل یک نرم افزار حرفه ای، ویژگی های قابل فعال سازی یا قابل غیرفعال سازی داشته باشد و تجربه کاربر را مطابق ترجیحات او تنظیم کند.

چرا این کار ارزشمند است؟

- همه کاربران نیازهای یکسانی ندارند.
- برخی می خواهند خروجی های اضافی مثل FAQ، گزارش، تحلیل یا خلاصه دریافت کنند، برخی نمی خواهند.
- امکان فعال / غیرفعال کردن ویژگی ها، GPT را مازولار، منعطف و کاربر محور می کند.

چگونه این کار را انجام می دهیم؟

با استفاده از Flipped Interaction Pattern.

در Prompt می گوییم:

- «همیشه ابتدا از کاربر بپرس که آیا می خواهد ویژگی X فعال شود یا نه.»
- «فقط پس از تعیین وضعیت ویژگی، وارد پاسخ گویی شو.»

مثال:

تولید خودکار FAQ Entry پس از هر پاسخ.

GPT ابتدا می پرسد: «آیا می خواهید قابلیت FAQ فعال شود؟»

اگر بله → پاسخ + تولید FAQ

اگر نه → فقط پاسخ

Introduction



Persona



Hallucination



Interactive



Ambiguity



Test



# GPT به عنوان شریک تصمیم‌گیری، نه پاسخ‌دهنده نهایی

هدف Custom GPT از حذف انسان از تصمیم‌گیری نیست، بلکه کمک به تصمیم‌گیری بهتر انسان است. مشکل رایج این است که GPT مستقیماً «جواب نهایی» را می‌دهد؛ در این حالت کاربر:

- سریع جواب را می‌بیند
- ادامه را نمی‌خواند
- فکر و بررسی را کنار می‌گذارد

طراحی درست این است که به جای جواب قطعی:

- اطلاعات و شواهد مرتبط (مثلًاً نقل قول از سیاست یا سند) ارائه شود
- نکات قابل تأمل و ملاحظات مطرح شود تا کاربر مجبور به فکر کردن شود

در این مدل، ما:

- هوش انسان را تقویت می‌کنیم، نه جایگزین
- تصمیم‌گیری را مشترک می‌کنیم (GPT + کاربر)
- اضافه کردن مسیر ارجاع به انسان (مثلًاً مدیر، HR، پشتیبانی) باعث می‌شود:
- GPT تنها نقطه تصمیم نباشد
- خطا کاهش پیدا کند

**Introduction****Persona****Hallucination****Interactive****Ambiguity****Test**



# کنترل خروجی Custom GPT با Template Pattern

یک روش حرفه‌ای در پرامپتنویسی است که به شما امکان می‌دهد:

- یک قالب کاملاً مشخص تعریف کنید
- بخش‌های قابل تغییر را به صورت جای خالی (Placeholder) تعیین کنید
- مدل را موظف کنید فقط همان قالب را حفظ کند
- و محتوا را دقیقاً در جای خالی‌ها قرار دهد
- این دقیقاً همان کاری است که با یک دستیار حرفه‌ای انجام می‌دهید:  
فرم می‌دهید، جای خالی تعریف می‌کنید، و انتظار دارد فرم دقیقاً همین‌گونه کامل شود.

PlaceHolders = جای خالی

Introduction



Persona



Hallucination



Interactive



Ambiguity



Test



# کنترل خروجی با Custom GPT

برای اینکه مدل به صورت کامل از قالب تبعیت کند، چهار اصل وجود دارد:

## 1. اعلام وجود قالب

به مدل می‌گوییم:

من یک قالب به تو می‌دهم.

این جمله مسیر ذهنی مدل را روشن می‌کند.

## 2. تعریف جای خالی‌ها

کلماتی که با حروف بزرگ نوشته می‌شوند نقش Placeholder دارند:

PRODUCT\_NAME

PRICE

CTA\_TEXT

مدل می‌فهمد این بخش‌ها باید با خروجی واقعی جایگزین شوند.

## 3. دستور پر کردن بخش‌ها

با یک جمله شفاف:

جای خالی‌ها را با محتوای مناسب پر کن.

## 4. اجبار به حفظ قالب

برای جلوگیری از تغییر ناخواسته ساختار:

قالب را بدون هیچ‌گونه تغییر حفظ کن.

این مرحله قلب Template Pattern است.

Introduction



Persona



Hallucination



Interactive



Ambiguity



Test

# Ambiguity management



# ضرورت تعریف مرزها در طراحی Custom GPT

رفتار کاربر و مدل همیشه قابل پیش‌بینی نیست، بنابراین تعریف مرزهای دقیق برای Custom GPT ضروری است.

- اگر مرزبندی وجود نداشته باشد، مدل ممکن است از مأموریت خود خارج شود، درخواست‌های نامناسب را اجرا کند یا به نتایج خطرناک و اشتباه برسد.
- باید مشخص کنیم که مدل چه کارهایی انجام می‌دهد و چه کارهایی انجام نمی‌دهد؛ وقتی کاربر درخواست مبهم یا خارج از دامنه می‌دهد، مدل چگونه باید واکنش نشان دهد.

مثال اول: کاربری درخواست ساخت تصویر گربه می‌دهد، در حالی که مدل اصلاً برای تولید تصویر طراحی نشده است. بدون مرز، مدل می‌تواند به مسیر اشتباه برود یا حتی مسیرهای سوءاستفاده را باز کند.

مثال دوم: درخواست «تضمین رعایت قوانین مالی»؛ مدل باید مرزی داشته باشد که هرگز ادعای تضمین نکند، چون چنین کاری ممکن نیست.

مثال سوم: درخواست مبهم «بهتر کردن خروجی قبلی». باید مرزی وجود داشته باشد که مدل تا وقتی هدف کاربر روشن نیست پاسخ ندهد و سؤال شفاف‌سازی بپرسد.

نتیجه‌گیری: مرزبندی یک بخش بنیادین در طراحی Custom GPT است؛ برای جلوگیری از سوءبرداشت، خروج از دامنه، تولید محتوای نادرست یا خطرناک، و تضمین اینکه مدل دقیقاً مطابق هدف سازنده رفتار کند.

Introduction



Persona



Hallucination



Interactive



Ambiguity



Test



# مدیریت ابهام در دانش‌پایه Custom GPT

هنگام استفاده از یک دانش‌پایه (Knowledge Base)، همیشه احتمال ابهام وجود دارد؛ زیرا هم متن سیاست‌ها ناقص است، هم پرسش‌های کاربران می‌توانند به چیزهایی اشاره کنند که در دانش‌پایه وجود ندارد.

ابهام از دو جهت ایجاد می‌شود:

- نقص یا گنج‌بودن خود سند دانش
- پرسش کاربر درباره چیزی که در سند وجود ندارد.

باید هنگام طراحی GPT مشخص کنیم که «ابهام» دقیقاً یعنی چه: وضعیتی که پاسخ نیازمند حدس‌زنی است و متن دانش‌پایه پاسخ صحیح ندهد.

برای کنترل این وضعیت، لازم است دستور صریح بدهیم که در صورت نبود پاسخ مستقیم:

- مدل نباید پاسخ را حدس بزند.
- مدل باید کاربر را یک پشتیبان انسانی ارجاع دهد. (مثلاً به واحد پشتیبانی)
- یا فقط پاسخ‌هایی ارائه کند که مستقیماً قابل استناد به متن هستند.

باید برای GPT «سوپاپ اطمینان» تعریف کنیم: وقتی اطلاعات کافی نیست، مدل بداند دقیقاً چه کار باید بکند؛ نه اینکه جواب بسازد یا تفسیر مبهم ارائه دهد.

بسته به هدف GPT، این سوپاپ می‌تواند:

- توقف کامل پاسخ و ارجاع
- تولید پاسخ با هشدار
- ارائه گزینه‌های جایگزین موجود در سیاست
- یا هر سازوکار دیگری باشد که با کارکرد GPT سازگار است.

Introduction



Persona



Hallucination



Interactive



Ambiguity



Test



# الگوی «اصلاح پرسش» برای رفع ابهام در Custom GPT

بسیاری از سؤالات کاربران مبهم هستند و Custom GPT باید راهی سیستماتیک برای مدیریت این ابهام داشته باشد. یکی از تکنیک‌های مؤثر، استفاده از Question Refinement Pattern است: مدلی که به صورت خودکار پرسش کاربر را تحلیل و نسخه‌ای واضح‌تر، دقیق‌تر و غیرمبهم از آن تولید می‌کند.

**مکانیزم کار:**

- کاربر سؤال اولیه را مطرح می‌کند.
- GPT نسخهٔ بهتر، دقیق‌تر و قابل پاسخ‌تری از سؤال را می‌سازد.
- GPT این نسخهٔ اصلاح‌شده را به کاربر نشان می‌دهد و می‌پرسد آیا تمایل دارد پاسخ بر اساس نسخهٔ بهبود یافته ارائه شود.

**مزیت کلیدی:** مدل بدون اینکه حدس بزند، نیت واقعی کاربر را روشن می‌کند؛ اگر اصلاح اشتباه باشد، خود کاربر آن را رد می‌کند و اطلاعات بیشتری در مورد نیت خود می‌دهد.

این الگوریسم را کاهش می‌دهد: مدل دیگر پاسخ اشتباه، خطرناک یا نامرتبط نمی‌دهد، زیرا قبل از پاسخ، سؤال را دقیق می‌کند.

نکنه مهم: اصلاح خودکار پرسش، هم به تولید پاسخ صحیح کمک می‌کند، هم به کاربر نشان می‌دهد که مدل قصد دارد ابهام را صفر کند. اگر کاربر بگوید «نه، این منظور من نبود»، ما اطلاعات دقیق‌تری درباره منظور او دریافت می‌کنیم؛ چیزی که در پاسخ مستقیم ممکن است اشتباه تفسیر شود.

Introduction



Persona



Hallucination



Interactive



Ambiguity



Test



# الگوی «اصلاح پرسش» برای رفع ابهام در Custom GPT

در بسیاری از موقعیت‌ها، نه دانش‌پایه کامل است و نه سؤال کاربر پاسخ شفاف دارد؛ اما همچنان باید به کاربر کمک کنیم بدون اینکه ریسک اشتباه، هزینه مالی یا پیامد حقوقی ایجاد شود.

چالش اصلی این است: چگونه هم مفید باشیم و به کاربر در حل مسئله کمک کنیم، و هم از پاسخ‌های پرخطر یا حدسی در شرایط مبهم اجتناب کنیم؟

راه حل مؤثر: استفاده از Alternative Approaches Pattern؛ یعنی وقتی پاسخ دقیق ممکن نیست، مدل به جای حدس زدن، راهکارهای مجاز، امن و مطابق سیاست ارائه می‌دهد.

سازوکار:

- مدل ابتدا سؤال را تحلیل می‌کند و تعیین می‌کند که آیا بر اساس دانش‌پایه می‌توان پاسخ قطعی داد.
- اگر پاسخ روشی است، به صورت عادی پاسخ می‌دهد.
- اگر پاسخ مبهم یا چندتفصیره است، مدل باید صریح بگوید: «پاسخ نامشخص است.»
- سپس به جای پاسخ دهنده مستقیم، سه راهکار جایگزین پیشنهاد می‌کند که کاملاً منطبق با سیاست و بدون ریسک باشند.

این الگو باعث می‌شود:

- مدل هرگز وارد محدوده خطرناک تفسیر اشتباه نشود.
- کاربر همچنان روش‌هایی برای حل مسئله داشته باشد.
- بین «کمک‌رسانی» و «رعایت مرزهای ایمنی و دانش» تعادل ایجاد شود.

نتایج کلیدی:

- کاهش ریسک حقوقی و مالی
- جلوگیری از پاسخ‌های حدسی
- افزایش کارآمدی و رضایت کاربر در شرایطی که پاسخ مستقیم ممکن نیست

Introduction



Persona



Hallucination



Interactive



Ambiguity



Test



# الگوی «راهکارهای جایگزین» برای پاسخگویی در شرایط مبهم

در بسیاری از موقعیت‌ها، نه دانش‌پایه کامل است و نه سؤال کاربر پاسخ شفاف دارد؛ اما همچنان باید به کاربر کمک کنیم بدون اینکه ریسک اشتباه، هزینه مالی یا پیامد حقوقی ایجاد شود.

چالش اصلی این است: چگونه هم مفید باشیم و به کاربر در حل مسئله کمک کنیم، و هم از پاسخ‌های پرخطر یا حدسی در شرایط مبهم اجتناب کنیم؟

راه حل مؤثر: استفاده از Alternative Approaches Pattern؛ یعنی وقتی پاسخ دقیق ممکن نیست، مدل به جای حدس زدن، راهکارهای مجاز، امن و مطابق سیاست ارائه می‌دهد.

سازوکار:

- مدل ابتدا سؤال را تحلیل می‌کند و تعیین می‌کند که آیا بر اساس دانش‌پایه می‌توان پاسخ قطعی داد.
- اگر پاسخ روشی است، به صورت عادی پاسخ می‌دهد.
- اگر پاسخ مبهم یا چندتفصیره است، مدل باید صریح بگوید: «پاسخ نامشخص است.»
- سپس به جای پاسخ دهنده مستقیم، سه راهکار جایگزین پیشنهاد می‌کند که کاملاً منطبق با سیاست و بدون ریسک باشند.

این الگو باعث می‌شود:

- مدل هرگز وارد محدوده خطرناک تفسیر اشتباه نشود.
- کاربر همچنان روش‌هایی برای حل مسئله داشته باشد.
- بین «کمک‌رسانی» و «رعایت مرزهای ایمنی و دانش» تعادل ایجاد شود.

نتایج کلیدی:

- کاهش ریسک حقوقی و مالی
- جلوگیری از پاسخ‌های حدسی
- افزایش کارآمدی و رضایت کاربر در شرایطی که پاسخ مستقیم ممکن نیست

Introduction



Persona



Hallucination



Interactive



Ambiguity



Test



# الگوی تحلیل پرسش Cognitive Verifier برای

بسیاری از پرسش‌ها تنها زمانی قابل پاسخ هستند که ابتدا به چند پرسش کوچک‌تر و دقیق‌تر شکسته شوند؛ زیرا سیاست‌ها شامل مفاهیم مبهمی مثل «مناسب و معقول» هستند که خودشان نیازمند تفسیر و پرسش‌های فرعی‌اند.

الگوی Cognitive Verifier کمک می‌کند مدل ابتدا پرسش کاربر را تحلیل کند، سه الی چهار پرسش جزئی‌تر بسازد، و سپس از پاسخ آن‌ها برای تولید یک پاسخ نهایی دقیق و قابل استفاده کند.

## سازوکار الگو:

- مدل پرسش کاربر را دریافت می‌کند.
- سؤال را از منظرهای مختلف خرد می‌کند و ۳-۴ سؤال فرعی حیاتی تولید می‌کند که پاسخ به آن‌ها، پاسخ اصلی را روشن می‌سازد.
- این پرسش‌های فرعی به کاربر ارائه می‌شود و از او پرسیده می‌شود آیا مایل است مدل با این فرایند ادامه دهد.
- در صورت تأیید، مدل سؤال‌های فرعی را تک‌به‌تک پاسخ می‌دهد و سپس آن پاسخ‌ها را ترکیب می‌کند تا نتیجه اصلی به دست بیاید.

## مزایا:

- ارتقای کیفیت استدلال مدل با تقسیم مسئله به اجزای کوچک‌تر
- جلوگیری از برداشت نادرست مدل از سؤال اصلی
- کمک به کاربر برای دیدن ابعاد پنهان سؤال، و تنظیم دقیق نیت خود
- ایجاد سازوکار شفاف برای رسیدن به پاسخ نهایی با استفاده از زیرپرسش‌های قابل استناد

Introduction



Persona



Hallucination



Interactive



Ambiguity



Test



## «تشخیص ابهام در نگاشت مفاهیم» برای جلوگیری از تفسیر اشتباه سیاست‌ها

یکی از رایج‌ترین انواع ابهام زمانی رخ می‌دهد که کاربر از مفهومی استفاده می‌کند که در سیاست وجود ندارد یا قابل‌نگاشت به چند مفهوم مختلف در دانش‌پایه است؛ این وضعیت می‌تواند به پاسخ‌های کاملاً متفاوت یا اشتباه منجر شود.

**مشکل اصلی:** مدل تلاش می‌کند اصطلاحات کاربر را به مفاهیم سیاست نگاشت کند؛ اما اگر چند نگاشت ممکن باشد، مدل ممکن است به اشتباه یکی را انتخاب کند و پاسخ نادرست بدهد.

**راحل مؤثر:** استفاده از Concept Mapping Ambiguity Pattern.

- مدل ابتدا سؤال را تحلیل می‌کند و بررسی می‌کند آیا مفاهیم مطرح شده در پرسش، به صورت قطعی و روشن در سیاست قابل‌نگاشت هستند یا خیر.
- اگر نگاشت واضح نباشد یا چند تفسیر ممکن وجود داشته باشد، مدل باید رفتار متفاوتی داشته باشد.

**سازوکار الگو:**

- اگر یک مفهوم در سیاست وجود ندارد یا چند نگاشت ممکن دارد، مدل باید:
- صریحاً اعلام کند: «پاسخ نامشخص است.»
- توضیح دهد که این مفهوم در سیاست وجود ندارد و قابل تفسیرهای مختلف است.
- به جای پاسخ مستقیم، چند پرسش جایگزین و قابل‌پاسخ ارائه دهد که مبتنی بر مفاهیم واضح و تعریف شده در سیاست باشند.

**مزایا:**

- جلوگیری از پاسخ‌های خطرناک یا نادرست ناشی از نگاشت اشتباه مفاهیم
- کمک به کاربر برای فهم اینکه چه چیزهایی در سیاست تعریف شده و چه چیزهایی نشده
- ارائه مسیرهای کمکی بدون پاسخدادن مستقیم به پرسش مبهم
- افزایش ایمنی و قابل‌اعتماد بودن GPT در مستندات سازمانی

Introduction



Persona



Hallucination



Interactive



Ambiguity



Test



# الگوی «حل تعارض میان چند پایگاه دانش» برای مدیریت و ابهام‌های چندمنبعی

هنگام استفاده از چند سند، چند سیاست یا چند منبع اطلاعاتی، تقریباً همیشه با تعارض‌ها و ابهام‌های بین سندی مواجه می‌شویم؛ به خصوص وقتی اسناد نسخه‌های مختلفی از یک سیاست هستند یا بخش‌هایی از هم را پوشش می‌دهند.

**مشکل اصلی:** اسناد می‌توانند اطلاعات متضاد یا ناقص ارائه دهند؛ در این حالت اگر GPT استراتژی مشخصی برای Conflict Resolution نداشته باشد، خروجی آن ناپایدار و غیرقابل اعتماد خواهد بود.

**راه حل:** تعریف صریح استراتژی حل تعارض در دستورالعمل GPT، تا مدل بداند هنگام برخورد با دو سند که اطلاعات متفاوت دارند دقیقاً چه کاری باید انجام دهد.

**دو استراتژی نمونه :**

**استراتژی ۱:** سند جدید اولویت دارد، اما سند قدیمی برای حوزه‌های پوشش‌داده‌نشده همچنان معتبر است.  
اگر سند تازه سیاستی را تغییر نداده، نسخه قدیمی همچنان اعمال می‌شود.

**استراتژی ۲:** سند جدید کامل و supersede است؛ هرچه در سند جدید نیست، دیگر معتبر نیست.  
سند تازه را نسخه نهایی و جامع می‌دانیم.

**همیت موضوع:**

- در بسیاری از سازمان‌ها سیاست‌ها با گذشت زمان تغییر می‌کنند و اسناد متعدد وجود دارد.
- بدون قواعد مشخص حل تعارض، GPT ممکن است رفتار متناقض یا نادرست داشته باشد.

**باید هنگام طراحی Custom GPT مشخص کنیم:**

کدام سند اولویت دارد؟

آیا اسناد قدیمی در نبود پوشش جدید معتبرند؟

آیا سند جدید همه‌چیز را جایگزین می‌کند یا فقط بخشی را؟

در صورت تعارض، مدل دقیقاً چه قوانینی را باید اجرا کند؟

Introduction



Persona



Hallucination



Interactive



Ambiguity



Test

# Testing and Evaluation of a CustomGPT



# چرا Custom GPT ضروری است؟

با هر تغییر در Custom GPT

دستورها (Instructions)

پایگاه دانش

Guardrail یا Persona

ممکن است رفتار غیرمنتظره نشان دهد.

مشکل رایج:

GPT در یک بخش بهتر می‌شود.

ولی در بخش دیگری پسرفت (Regression) می‌کند.

بررسی موردی و حسّی (ad-hoc):

قابل اعتماد نیست

باعث می‌شود بعضی ضعف‌ها اصلاً دیده نشوند.

Benchmark باعث می‌شود:

کیفیت واقعی قابل مشاهده شود.

بدانیم GPT «واقعاً» در چیزهای مهم خوب عمل می‌کند.

Introduction



Persona



Hallucination



Interactive



Ambiguity



Test



# عملی چه اجزایی Benchmark دارد؟

سenarioهای تست:

- کاربردهای اصلی مورد انتظار
- Edge case ها و درخواست های غیرمنتظره

انتظار از خروجی (Expected Output):

- توصیف کیفی پاسخ خوب
- یا نمونه خروجی مطلوب

ارزیابی (Rubric):

- معیارهای مشخص برای قضاوت
- چه چیزی امتیاز را کم یا زیاد می کند

نمره دهنده ساده:

- مثلًا 1 تا 10
- برای مقایسه نسخه ها مختلف در طول زمان

Introduction



Persona



Hallucination



Interactive



Ambiguity



Test



# چیست؟ What If

What If یک الگوی تفکر سناریومحور است که برای بررسی رفتار یک سیستم، تصمیم، یا مدل در شرایطی به کار می‌رود که:

- غیرایده‌آل هستند،
- به طور مستقیم در مشخصات اولیه نیامده‌اند،
- اما در دنیای واقعی محتمل‌اند.
- 

به صورت خلاصه:

«اگر شرایط از حالت نرمال خارج شد، سیستم چه می‌کند؟»

Introduction



Persona



Hallucination



Interactive



Ambiguity



Test



Introduction



Persona



Hallucination



Interactive



Ambiguity



Test

# سناریو ۱ : GPT پشتیبانی مشتری برای یک شرکت مخابراتی

سناریوهای What If برای آزمون:

- اگر مشتری نارضایتی خود را به صورت غیرمستقیم بیان کند چه می‌شود؟  
- آزمون توانایی GPT در تشخیص زبان منفعل نشان‌دهنده نارضایتی و پاسخ‌دادن با همدلی و تکنیک‌های کاهش تنش.
  
- اگر مشتری از اصطلاحات فنی به صورت نادرست استفاده کند چه می‌شود؟  
- آزمون اینکه آیا GPT می‌تواند به نرمی مشتری را اصلاح کرده و اطلاعات درست را بدون ایجاد سردگمی یا رنجش ارائه دهد.
  
- اگر مشتری درباره خدمتی یا محصولی سؤال کند که وجود ندارد چه می‌شود؟  
- آزمون توانایی GPT در هدایت مشتری به گزینه‌های موجود و مدیریت انتظارات.



# سناریو 2: GPT به عنوان دستیار مشاوره مالی

Introduction



Persona



Hallucination



Interactive



Ambiguity



Test

سناریوهای What If برای آزمون:

- اگر کاربر درباره یک اقدام سرمایه‌گذاری غیرقانونی یا غیراخلاقی درخواست راهنمایی کند چه می‌شود؟  
- آزمون پایبندی GPT به استانداردهای قانونی و اخلاقی و توانایی آن در رد ارائه کمک در چنین مواردی.

- اگر کاربر اطلاعات ناکافی یا نادرستی درباره وضعیت مالی خود ارائه دهد چه می‌شود؟  
- آزمون نحوه برخورد GPT با نیاز به اطلاعات کامل و دقیق برای ارائه مشاوره قابل اتکا، احتمالاً از طریق طرح پرسش‌های تکمیلی.

- اگر کاربر پیش‌بینی حرکت بازار را بخواهد چه می‌شود؟  
- آزمون توانایی GPT در مدیریت انتظارات و انتقال غیرقابل پیش‌بینی بودن ذاتی بازارهای مالی، در عین ارائه توصیه‌های کلی مبنی بر داده‌های تاریخی.



## Custom GPT برای محافظت از (Adversarial Testing)

اگر Custom GPT قرار است در اختیار کاربران غیرقابل اعتماد یا عمومی قرار گیرد، حتماً باید آزمون خصمانه انجام دهیم؛ زیرا کاربران می‌توانند تلاش کنند مدل را دور بزنند، گمراх کنند، یا وادار کنند به گفتن چیزهایی که به سازمان آسیب می‌زند.

آزمون خصمانه یعنی: تلاش عمدى برای شکستن گاردrielها، دستکاری مدل، تولید خروجی‌های نامناسب، یا وادار کردن مدل به انجام کاری که نباید انجام دهد؛ دقیقاً همان کاری که یک مهاجم انجام می‌دهد.

یکی از خطرات اصلی: بیان روایت‌ها از زبان شما یا سازمان شما. کاربر می‌تواند روایت جعلی بسازد، سپس بخشی از خروجی مدل را بدون زمینه منتشر کند و ادعا کند «GPT سازمان این را گفته است».

خطر دوم: فعال بودن قابلیت‌هایی که نباید فعال باشند، مثل تولید تصویر. با چند دور دستکاری، مدل تصاویر تولید می‌کند؛ کاری که نباید در یک GPT سیاست‌های سفر سازمان انجام دهد.

### درس کلیدی:

- اگر یک قابلیت (مثل تولید تصویر، دسترسی API، اجرای کد) نیاز واقعی ندارد، باید از ابتدا غیرفعال شود.
- اگر فعال است، باید قوانین سختگیرانه و پاسخ‌های محافظه‌کارانه تعریف کنیم.

### چرا این مهم است؟

- مدل می‌تواند ناخواسته خروجی‌هایی تولید کند که سازمان را در معرض خطر حقوقی، اعتباری یا رسانه‌ای قرار می‌دهد.
- کاربر خصمانه فقط بخش آسیب‌زا را منتشر می‌کند، نه کل مکالمه را.

### باید هنگام طراحی GPT:

- گاردriel‌های قوی تعریف کنیم.
- قابلیت‌های غیرضروری را حذف کنیم.
- سناریوهای حمله را شناسایی و شبیه‌سازی کنیم.
- بررسی کنیم آیا مدل می‌تواند وادار شود از چارچوب خارج شود یا خیر.

### Introduction



### Persona



### Hallucination



### Interactive



### Ambiguity



### Test



# برای ارزیابی خروجی (Rubric)

روبیک ارزیابی پاسخ‌های GenAI می‌تواند شامل عوامل کلیدی زیر باشد:

کیفیت استدلال:

- درستی پاسخ‌ها
- انسجام منطقی
- شواهد درک مفاهیم پیچیده
- اثربخشی حل مسئله

لحن و سبک:

- تناسب با زمینه و لحن کاربر
- سازگاری با سبک مکالمه مورد انتظار

کامل بودن:

- پاسخ دادن به همه بخش‌های پرسش‌های چندوجهی
- ارائه جزئیات کافی در صورت نیاز

دقت:

- صحت واقع‌محور
- پایبندی به دستورالعمل‌ها یا راهنمایی داده شده

مرتبه بودن:

- ارتباط مستقیم پاسخ با پرسش مطرح شده
- پرهیز از اطلاعات حاشیه‌ای یا نامرتبط

ایمنی و انطباق:

- عدم تولید محتوای مضر
- بی‌طرفی در خروجی
- تناسب فرهنگی برای کاربران هدف
- احترام به حریم خصوصی و حفاظت از داده‌ها
- پایبندی به استانداردهای قانونی و اخلاقی

Introduction



Persona



Hallucination



Interactive



Ambiguity



Test



# ارزیابی ویژگی‌های مکالمه‌ای در چند پیام

انسجام (Coherence):

- ارتباط زمینه‌ای: پیام‌ها به زمینه قبلی مرتبط باشند.
- جریان منطقی: پیام‌ها به صورت منطقی بر یکدیگر بنا شوند.
- شفافیت ارجاع: موضوعات قبلی به درستی و باوضوح ارجاع داده شوند.

پیوستگی (Continuity):

- حفظ موضوع: پایبندی به موضوع اصلی در چند پیام.
- نرمی انتقال: جایه‌جایی نرم بین موضوعات در مکالمه.
- یادآوری تعاملات قبلی: استفاده و ارجاع به اطلاعات تبادلات پیشین.

پاسخ‌گویی (Responsiveness):

- به موقع بودن: پاسخ‌های سریع که ریتم مکالمه طبیعی را حفظ کند.
- صراحة: هر پاسخ مستقیماً به نکات پیام قبلی پردازد.
- تأیید و اذعان: نشانه‌هایی که نشان دهد هوش مصنوعی پیام کاربر را درک کرده یا با آن موافق است.

کیفیت تعامل:

- درگیرسازی: حفظ علاقه کاربر از طریق گفت‌وگوی تعاملی.
- همدلی و آگاهی عاطفی: تشخیص و پاسخ مناسب به نشانه‌های احساسی.
- شخصی‌سازی: تطبیق مکالمه بر اساس تعاملات و ترجیحات پیشین کاربر.

مدیریت مکالمه:

- بازیابی خطأ: مدیریت و اصلاح سوءبرداشت‌ها.
- ادب و آداب: رعایت هنجارهای ارتباط محترمانه.
- رفع ابهام: تلاش برای شفاف‌سازی عدم قطعیت‌ها یا ابهامات گفت‌وگو.

تکامل (Evolution):

- پیشرفت: پیشبرد تم‌ها یا روایت‌ها با پیش‌روی مکالمه.
- یادگیری و تطبیق: اصلاح گفت‌وگو بر اساس تاریخچه مکالمه و بازخورد کاربر.
- جمع‌بندی و پیگیری: پایان‌دهی مناسب مکالمه و فراهم کردن زمینه تماس‌های بعدی.

Introduction



Persona



Hallucination



Interactive



Ambiguity



Test



# Test Case تولیدکننده GPT

- طراحی Test Case خوب سخت است
- انسان‌ها همهٔ حالت‌ها، ریسک‌ها و سوءاستفاده‌ها را نمی‌بینند
- تست ناقص → ضعف پنهان → شکست در دنیای واقعی

راه حل پیشنهادی:

- ساخت یک Custom GPT مخصوص تولید Test Case
- ریسک پایین: خروجی فقط «ایدهٔ تست» است، نه تصمیم عملی

نحوه کار این GPT:

- با سوال‌های مرحله‌ای می‌فهمد GPT اصلی چه کاری می‌کند.
- از یک سند «Design Considerations» الهام می‌گیرد.

Test Case های متنوع تولید می‌کند شامل:

- هدف تست
- پرامپت کاربر
- انتظار از خروجی
- ارزیابی Rubric

مزیت کلیدی:

- پوشش سناریوهای واقعی، مبهم و پر ریسک
- امکان تمرکز روی safety، adversarial، compliance، هزینه،
- قابل خروجی‌گرفتن به Benchmark برای مدیریت CSV / Excel

با GPT تولیدکننده Test Case از حدس انسانی به فرآیندی سیستماتیک تبدیل می‌شود.

Introduction



Persona



Hallucination



Interactive



Ambiguity



Test



Introduction



Persona



Hallucination



Interactive



Ambiguity



Test

# تنوع در موارد آزمون

برای پوشش طیف تعاملات و چالش‌های کاربر، موارد آزمون باید در چند بعد بسته به اهداف متنوع باشند:

نوع وظیفه/پرسش:

- پرسش‌های واقع‌محور (مثلاً پرسش‌های ساده درباره اطلاعات شناخته‌شده)
- وظایف استدلالی (مثلاً معماها یا پرسش‌های حل مسئله)
- وظایف خلاقانه (مثلاً تولید داستان یا ایده)
- وظایف مبتنی بر دستورالعمل (مثلاً راهنمایی مرحله‌به‌مرحله)

ویژگی‌های کاربر:

- سطح سواد (مثلاً پایه، متوسط، پیشرفته)
- دانش حوزه‌ای (مثلاً غیرمتخصص، علاقه‌مند، متخصص)
- زبان و گویش‌ها (مثلاً تنوع‌های زبانی، گویشوران غیر بومی)
- ویژگی‌های جمعیت‌شناختی (مثلاً سن، پیش‌زمینه فرهنگی)

پیچیدگی ورودی:

- طول ورودی (مثلاً یک جمله، پاراگراف، گفت‌وگوی چندنوبتی)
- شفافیت زمینه (با یا بدون زمینه کافی)
- ابهام و کلی‌گویی در پرسش‌ها

لحن یا احساسات ورودی

- ورودی‌های خصم‌مانه (Adversarial)
- پرسش‌های عامدانه گمراه‌کننده یا فریبنده
- تلاش برای القای پاسخ‌های جانبدارانه یا نامناسب
- ورودی‌هایی با هدف نقض حریم خصوصی یا استانداردهای امنیتی