Use of AI/Chat Models in the 401(k) Industry – A Practical Guide

ERISApedia

1

Presented by:

- Samantha Grimes

- Tim McCutcheon

ERISApedia

2

Housekeeping Items

- All attendee lines are muted. Please use the Q&A button on zoom toolbar to ask questions. Please do not use chat. We will get to as many questions as possible during the webinar but will follow-up personally to any unanswered questions.
- Click on "Live Transcript/Closed Captioning" to see a written transcript of what is being said.

ERISApedia

3

Agenda

- Introduction to AI Models
- Prompt Engineering
- How ERISApedia uses AI models to help craft communications with prospects and clients.

ERISApedia

4

AI/Chat - What is it?

- Chat GPT
  - A **large language model** designed to follow user intentions using techniques developed through alignment research. These models are trained with humans in the loop.
  - ERISApedia: gpt-4 and gpt-3.5-turbo
- PaLM 2/Bard
  - PaLM 2 is Google's next generation **large language model** that builds on Google's legacy of breakthrough research in machine learning and responsible AI.
  - ERISApedia: chat-bison

ERISApedia

5

Large Language Models

- What is a large language model? We asked Chat GPT-3.5:
  - A large language model is a type of artificial intelligence system that is designed to understand and generate human-like text. It uses machine learning techniques to analyze vast amounts of text data and learn patterns, structures, and relationships within the data. These models are typically trained on massive datasets that contain diverse and representative samples of human language.
  - Large language models, such as OpenAI's GPT-3 (Generative Pre-trained Transformer 3), consist of deep neural networks with multiple layers and millions or even billions of parameters. These models can generate coherent and contextually relevant text based on a given prompt or question.
- Also known as Generative AI

ERISApedia

6

How Do They Work?

- User enters a prompt in plain language, the prompt gets converted to numeric tokens, [magic happens], the model responds by converting response tokens to plain language.
  - Chat GPT: https://chat.openai.com/
  - Bard: https://bard.google.com/.
- The quality of the prompt determines the quality of the response. Garbage in - garbage out.
- The art of writing effective prompts is known as prompt engineering.

ERISApedia

7

Types of Prompts

- Question Answering
- Chat Simulation
- Content Generation
- Summarization
- Data Extraction

ERISApedia

8

Prompt Engineering

- Prompt Framing
  - You are acting as a third-grade teacher/college professor. Explain gravity.
  - You are acting as an experienced financial advisor specializing in 401(k) plans.
- X-Shot Learning
  - Provide zero to many examples of what you are expecting the model to do:
    - Prompt: I don't like product x Response: Unfavorable
    - Prompt: I think ice cream tastes great
    - Response [enter]
- Prompt Length and Specificity
  - Talk to the model like it was smart teenager
  - Use consistent language
  - Delimit data to be analyzed

ERISApedia

9

Summarization Demo

- Version One:
- What are the vesting rules in the following paragraph:

ERISApedia

10

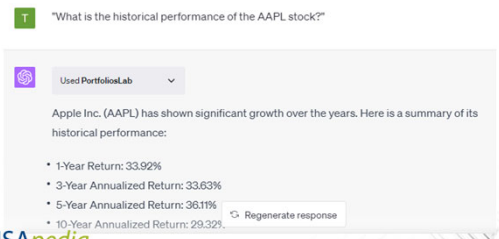
Summarization Demo

- Version Two:
- You are acting as a legal assistant to summarize the vesting schedules in a description of a 401(k) plan that is delimited by 3 back ticks. ***There are generally three types of vesting schedules, the first applies to employee contributions, the second applies to employer/company non-elective/profit-sharing contributions and third applies to employer/company matching contributions. Sometimes the description of a 401(k) plan will contain information about vesting schedules that are currently effective and vesting schedules that were effective in the past***. Your summary will only contain information about vesting schedules that are currently effective. Your summary will contain three bullet points each of which will describe one of the three types of vesting schedules.

ERISApedia

11

Plug-ins



ERISApedia

12

Search

- Bing:
  - What large language model does Bing use?
  - Bing uses a new next-generation OpenAI large language model that is more powerful than ChatGPT and customized specifically for search. The new OpenAI large language model that powers Bing has been optimized for search. It's built on learning from ChatGPT and GPT-3.52.
- Google:
  - Bueller? Bueller? Bueller?
  - https://labs.withgoogle.com/sge/ (Search Labs - limited signup)

ERISApedia

13

How AI/Chat Models Are Used in Third Party Products Like ERISApedia

- There are two ways to access the chat models.
  - Human interface
  - Application Programming Interface (API) - The API works in a similar manner as the human interface except a third-party server sends a prompt to OpenAI or Google servers and the response is sent back to the third-party servers.
- Third party products like ERISApedia use the API.
- Large organizations may train their own models
  - Zoom
  - BloombergGPT

ERISApedia

14

LLM Limitations

- Token limit
  - 8k – 32k
- Randomness
- Hallucinations
  - Hallucination in AI refers to the generation of outputs that may sound plausible but are either factually incorrect or unrelated to the given context.

ERISApedia

15

ERISApedia Process

- Determine the nature of your campaign
- Do a search for plans meeting your campaign criteria
- **Select report criteria/KPIs/Send communication - New!!**
- Win business

ERISApedia

16

ERISApedia Communications

- Highly-engineered prompt delivered via API
- Basis of prompt is a text-only version of the ERISApedia Benchmark Report
  - Pick characteristics: Four Types, Good Only/Bad Only/Good and Bad
- Maximum control over content
  - Model
  - Open/closing statements
  - Randomness

ERISApedia

17

ERISApedia Benchmark Sample KPI

Google LLC
Google LLC 401(k) Savings Plan
Detail Operational Benchmark Results

Average Participant Contribution

Average Participant Contribution is determined by dividing the total participant contributions by the total active participants in the plan at the end of the year. A red gauge means that the plan rates unfavorably (low percentile score) and a green gauge means that the plan rates favorably (high percentile score) versus the peer group.

A high percentile score generally indicates employee satisfaction with the retirement plan.

ERISApedia

18

ERISApedia Benchmark Report

- Nine Operational Benchmark Metrics/KPIs
  - High score favorable/good
- Plan Alerts – good things like 404(c).
  - Presence favorable/good
  - Choice of over some items
- Status Alerts – generally things that are going wrong.
  - Absence favorable/good
  - Choice of over some items
- Feature Prevalence

ERISApedia

19

Operational Benchmarks/KPIs

Benchmark Scores - Items selected: 9

Select All Deselect All

- Average Participant Contribution: ☑
- Average Employer Contribution: ☑
- Participation Rate: ☑
- Average Account Balance: ☑
- Average Benefit Distributions/Participant: ☑
- Corrective Distributions/Participant: ☑
- Percentage of Terminated Participants: ☑
- Investment Return: ☑
- Administrative Expense/Participant: ☑

ERISApedia

20

Plan Alerts

Plan Alerts - Items selected: 4

Select All Deselect All

- Cross Tested Plan: ☐
- Complies with DOL sec. 404(c) : ☐
- Participant Self Direction: ☑
- Matching Contributions: ☐
- Plan Allows Brokerage Accounts: ☑
- Auto Enrollment: ☑
- Default Investment Accounts: ☑

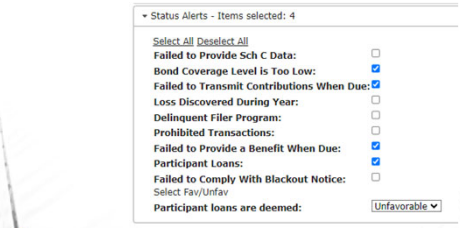
Select Fav/Unfav

- Brokerage accounts are deemed: Favorable
- Auto enrollment is deemed: Favorable

ERISApedia

21

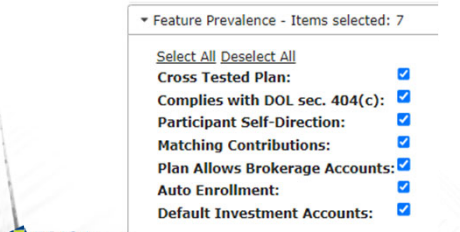
Status Alerts



ERISApedia

22

Feature Prevalence



ERISApedia

23

Types of Communications

- Experienced 401(k) Financial Advisor
  - Email
  - Phone script
- Custom framing/role
  - Sample: You are a compliance expert specializing in 401(k) retirement plans.
  - Email
  - Phone script

ERISApedia

24

Live Demo

- Formulate campaign
  - Work with plans that have low Participation Rate and low Average Account Balance.
- Select plans
  - Select Work plans that have low Participation Rate and low Average Account Balance
  - 401(k) $5M to $10M
- Select criteria
  - Only Benchmark scores
- Send communication

ERISApedia

25

Questions

- Samantha will follow-up personally with any questions we didn't get to.
- If you would like a recording of this webinar

Please contact Samantha at at:

- sgrimes@erisapedia.com
- 405-515-9022

26