

[Open Peer Review on Qeios](#)

# How Factually Accurate is GPT-3? A Focused Case Study on Helping Malaysia's B40s Through e-Commerce

Nabila Ameerah Zainal Abidin<sup>1</sup>, Khairul Khaizi Mohd Shariff<sup>1</sup>, Ihsan Mohd Yassin<sup>1</sup>, Azlee Zabidi<sup>2</sup>, Azlinda Saadon<sup>1</sup>, Nooritawati Md Tahir<sup>1</sup>, Abdul Rahim Ridzuan<sup>1</sup>, Megat Syahirul Amin Megat Ali<sup>1</sup>

<sup>1</sup> Universiti Teknologi Mara

<sup>2</sup> University Malaysia Pahang

**Funding:** The authors would like to thank the Ministry of Education, Malaysia for funding this research through the Long-Term Research Grant (LRGS) grant (Grant number 600-RMC/LRGS 5/3 (001/2020), Project Code: 251905210001).

**Potential competing interests:** No potential competing interests to declare.

## Abstract

GPT-3 (Generative Pre-trained Transformer 3) is an advanced natural language processing model utilizing unsupervised learning to generate sophisticated human-like text. GPT-3 has been lauded for its potential to revolutionize the field of natural language processing, with its capacity to generate a variety of text with a high degree of fluency and accuracy. We examine the ability of GPT-3's to produce text related to a focused subject matter: alleviating poverty in Malaysia through e-Commerce. We especially examine GPT-3's ability to produce factual responses within this narrow context. It was discovered that GPT-3 could produce plausible statements, albeit some of them being factually debatable or incorrect due to how its training data was sourced. We also discuss how GPT-3 could be used unscrupulously to either produce academic-sounding responses that appear to be a product of research, but possibly untrue or inaccurate and discuss its potential ramifications (such as propaganda and disinformation). We end the paper with some suggestions to the brilliant team at OpenAI to further improve GPT-3 for the advancement of humankind.

**Nabila Ameerah Zainal Abidin<sup>1</sup>, Khairul Khaizi Mohd Shariff<sup>1</sup>, Ihsan Mohd Yassin<sup>1,\*</sup>, Azlee Zabidi<sup>2</sup>, Azlinda Saadon<sup>3</sup>, Nooritawati Md Tahir<sup>4</sup>, Abdul Rahim Ridzuan<sup>4</sup>, and Megat Syahirul Amin Megat Ali<sup>1</sup>**

<sup>1</sup>*Microwave Research Institute, Universiti Teknologi Mara, Malaysia*

<sup>2</sup>*Fakulti Komputeran, Universiti Malaysia Pahang, Malaysia*

<sup>3</sup>*College of Engineering, Universiti Teknologi Mara, Malaysia*

<sup>4</sup>*Institute for Big Data Analytics and Artificial Intelligence, Universiti Teknologi Mara, Malaysia*

**Keywords:** Large Language Model (LLM), Generative Pre-trained Transformer 3 (GPT-3), Bottom 40 (B40), Natural Language Processing (NLP), E-Commerce.

## 1. Introduction

The impressive abilities of large language models (LLMs) via few-shot prompting have been demonstrated recently, with commercialized APIs such as GPT-3 further increasing their use in real-world language applications [1]. GPT-3 is a Natural Language Processing (NLP) model developed by OpenAI [2] that uses machine learning to generate sophisticated human-like text [3] using an autoregressive language transformer model. This model structure allows it to capture long-term dependencies in the training data, allowing the model to predict the next word in a sentence based on context and previously generated words [3].

The language model was constructed with an astounding 175 billion parameters trained on a massive 45 TB of data from various curated and uncurated sources, including books, articles, forums, and webpages [3], [4]. The GPT-3 training dataset is composed of 8 million web documents (collected using CommonCrawl and WebText), 40,000 books (from the Project Gutenberg digital library), and a wide range of other uncurated sources, including articles, question-answer pairs, and conversational dialogues [3]. The remaining sources of data for the GPT-3 training dataset are articles from Reddit [5], question answer pairs from Quora [6], and conversational dialogues from the Ubuntu Dialogue Corpus [7] covering a broad range of topics and interests. The extensive dataset GPT-3 was trained on suggests that it has learnt somewhat rudimentary form of knowledge source from the internet [3], [8].

The introduction of GPT-3 to the public has been met with excitement as the technology was hailed as able to generate human-like text in unprecedented manner [9]. However, reference [9] notes the limitations of the technology, identifying these flaws: (1) The output may lack semantic coherence as GPT-3 is required to write longer texts, (2) the output may be biased as a result from its training data, and even produce possibly radical themes [10], its responses **may not always be truthful, though appearing to be plausible** [11].

We believe that the above findings, in part, is because of the introduction of uncurated data in training the knowledge base of GPT-3. While this information is invaluable for GPT-3 to understand a broad set of topics and nuances of conversational discussions, the source makes its factual accuracy debatable. Importantly, the factuality of these uncurated sources needs to be examined in order to make GPT-3 useful for academic research that obviously discovers and evaluates facts using the scientific method.

In this paper, we attempted to test the factuality of GPT-3 on a focused case study about the Bottom-40's income earners in Malaysia, and how establishing an e-commerce business can assist their livelihood. The authors happened to obtain a research grant that focuses on this topic, and we consider this to be an ideal case study for GPT-3 because we wanted to establish whether GPT-3 can generate factually correct content over a relatively narrow study scope. This case study attempts to assess GPT-3's ability to write on topics within a global context, such as wealth inequality, and e-commerce. The information centered around these themes should be generally abundant, and thus GPT-3 should be able to generalize them well with a high degree of plausibility (or perhaps factuality). Additionally, the case study also contains elements within a narrower context, i.e. elements specific to Malaysia. For example, the term B40 is commonly used in Malaysia, but relatively unused/unknown outside Malaysia. Additionally, we tested GPT-3's performance in generating

texts about particularly sensitive topics, such as economic inequality in Malaysia.

The remainder of this paper is organized as follows: Section 2 describes recent relevant works, while Section 3 describes the research methodology. The results and discussion section is presented in Section 4. Concluding remarks are presented in Section 5. Finally, some suggestions on the future direction is GPT-3 is suggested in Section 6 to address some of the issues presented in this paper.

## 2. Recent Relevant Works

This section reviews recent works focused on improving LLMs (including GPT-3) reliability and factuality.

References [12], [13] found that GPT-3 reported factual data in Closed-Book Question Answering (CBQA) tests, while [1] found that GPT-3 was found to be more reliable than smaller-scale supervised LLMs.

Reference [14], [15] discovered that **GPT-3 is reliable and adaptable for generating high-quality news and user review summaries**. However, several articles reported that the flaw of GPT-3 is that it **cannot differentiate truth from falsehood** as its fundamental purpose is to generate text and not evaluate its factual validity<sup>[11]</sup>. Interestingly, other researchers [1], [13], [14] claim good factual accuracy of GPT-3.

Researchers in [12] introduced a contextual framework that utilizes a coarse-to-fine approach to extract relevant knowledge and answer a question without access to any external knowledge to improve LLM's ability to answer Closed Book Question-Answer (CBQA) tasks. The framework employed a pretrained language model to generate a related context for a given question, and to prompt the same language model for answer prediction based on the generated context and the question. On three benchmarks, it has been demonstrated that this method **outperforms state-of-the-art CBQA methods** (exact matching 68.6% vs. 55.3%) and is comparable to open-book methods that utilize external knowledge sources (68.6% vs. 68.0%). The proposed approach enabled better exploitation of the stored knowledge in pretrained LLMs without the need for extra learnable parameters or fine-tuning, and thus has the potential to be implemented as hybrid models that integrate pretrained LLMs with external knowledge.

Similarly, in [13], the authors evaluated GPT-3 performance on CBQA, a task that requires a significant amount of factual knowledge. By confining GPT-3 to the closed-book setting, external knowledge bases could not be retrieved and only the knowledge stored in the model's parameters are available. Analysis of GPT-3's performance on different types of questions not only **revealed the amount of factual knowledge stored** in GPT-3, but also offered a better understanding of its strengths and weaknesses in terms of reasoning capabilities. The authors found that GPT-3 performed well on factual questions but relatively underperformed on multi-hop, common sense and time-sensitive questions.

Researchers in [1] investigated the application of four main facets to improve GPT-3's reliability: generalizability, social biases, calibration, and factuality. An important contribution of the paper was establishing simple and effective prompts that improve GPT-3's reliability, as it generalizes out-of-distribution, balances demographic distribution and uses natural language instructions to reduce social biases, calibrates output probabilities, and updates the LLM's factual knowledge and reasoning chains. With the appropriate prompts, GPT-3 was found to be **more reliable than smaller-scale**

**supervised models** on all these facets. This systematic empirical study has not only provided new insights on the reliability of prompting LLMs, but more importantly, the prompting strategies developed can help practitioners use LLMs like GPT-3 more reliably.

Recent years have seen the leading paradigm of text summarization research being the fine-tuning of pre-trained models on domain-specific datasets. The impact of GPT-3 and two other LLMs on this was studied by [14], with focus on news summarization. The researchers found that GPT-3 summaries were of **extremely high-quality and adaptable** to different summarization settings, with a clear preference for them being observed in results from the human study across both datasets and styles. The use of zero-shot models like GPT-3 Da-Vinci 2 model for driving research in a more top-down way, as well as its improvements upon prior keyword-focused summarization systems that were trained on artificially adapted training data, was also demonstrated.

Researchers in [15] explored several pipeline methods for using GPT-3 to summarize user reviews without training, such as recursive summarization and supervised clustering/extraction. The models achieved strong performance on two datasets of hotel and product reviews. GPT-3 was found to produce **high-quality summaries**, although errors included subtle issues of **balancing contradictory opinions** and **generalizing specific claims**. Metrics were introduced to measure the factuality, faithfulness, and genericity of summaries, which showed that GPT-3 was more successful for shorter input reviews and hierarchical summarization led to generalities as the input size increased.

A study by [16] evaluated the ability of human subjects to discern between authentic and GPT-3 generated synthetic information structured as tweets. As part of the survey, for each subject category, the research included 20 tweets, five each of "organic true", "synthetic true", "organic false", and "synthetic false" varieties. The results indicated that in several controversial topics, the **subjects have difficulty distinguishing between tweets composed by GPT-3 and those written by human users**. The authors concluded that suggest that GPT-3 possesses a **double-edged capability**; on one hand, it can produce accurate information that is more easily understood than that produced by humans; on the other hand, it can **generate more persuasive disinformation**. Considering these results, the authors reflect on the risks associated with Artificial Intelligence (AI) for **disseminating disinformation**.

Reference [11] tested the restrictions of LLMs like GPT-3 by leveraging concepts such as compression, priming, distributional semantics, and semantic webs. The researchers discovered that GPT-3 could carry out a variety of semantic tasks. However, several of its answers often appear senseless because it was not designed to fact-check its answers. It was argued that these types of models were unable to be constrained to just yield accurate continuation and that they strategize to be **believable instead of truthful** to optimize their objective function. Furthermore, the researchers warned that extensive utilization of language generators as text-writing tools could result in a **lasting contamination of the informational environment** with a great deal of **very believable but frequently inaccurate** texts. Interestingly, another study [10] also affirmed the conclusions by [11] that **information contamination** may be a distinct possibility as the use of LLMs becomes prevalent – even to the point of it potentially being used to **propagate radicalization** of individuals.

### 3. Methodology

The authors basically wrote several questions to the GPT-3 Advanced Programming Interface (API) in the OpenAI playground. The questions were structured to test GPT-3's "knowledge" on several aspects:

1. Itself
2. Is it an information / knowledge aggregator / integrator / consensus?
3. The source of its knowledge base.
4. A case study on wealth inequality in Malaysia and how e-commerce can help the Bottom 40s (B40s) in Malaysia.

We then analyzed its answers to our queries. The questions and GPT-3's verbatim responses are shown in Section 4. The GPT-3 Playground has several settings that can be adjusted. They were set based on trial-and-error and viewing the GPT-3's responses on several questions. The parameters and their values are presented in Table 1.

**Table 1.** GPT-3 parameters used in the experiment.

Parameter	Value	Description
<b>Model</b>	davinci-003	The latest GPT-3 model from OpenAI.
<b>Temperature</b>	0.7	Controls the randomness of the output. A lower value results in a more deterministic and repetitive output.
<b>Maximum Length</b>	3,000	These tokens get used when we ask questions to the GPT-3 engine. Longer token lengths generally produce longer responses.
<b>TopP</b>	1	This is a parameter to control the diversity of GPT-3's responses via nucleus sampling.
<b>Frequency Penalty</b>	0.3	This setting penalizes the responses based on the text that has already appeared in the responses. A high setting reduces the tendency of GPT-3 to repeat certain texts.
<b>Presence Penalty</b>	0.3	This setting penalizes GPT-3 when a certain topic is repeated. An increase in this setting would improve the likelihood of GPT-3 to talk about new topics.
<b>Best of</b>	1	If set to 1, it simultaneously generates many responses and chooses what it considers to be the best response.

### 4. Results & Discussions

In this section, we will look at the responses of GPT-3 about several topics:

- What does the GPT-3 say about itself? (Section 4.1).
- GPT-3's Knowledge Base (Section 4.2).
- Is GPT-3's Information Factually Accurate? (Section 4.3).
- Case Study: Alleviating Poverty in Malaysia using e-Commerce according to GPT-3 (Section 4.4).

#### 4.1. What does the GPT-3 say about itself?

In this section, we asked GPT-3 about itself to establish its origins, knowledge base, and whether the model can give factual answers to questions given. Facts, of course, are fundamentally important in academic reporting. Additionally,

based on these questions, we attempt to uncover the source of GPT-3's knowledge, which gives us an insight to how it answers queries presented to it. Table 2 shows the questions and responses given by GPT-3. Due to the way the questions are structured, some information appears to be redundant. However, for the sake of brevity, the authors have chosen to include the answers verbatim. Referring to Table 2, we started with Question 1, a relatively straightforward question asking GPT-3 to describe itself in its own words. These answers appear to be consistent with general knowledge about GPT-3 and academic sources.

**Table 2.** Responses by GPT-3 about itself and whether it an information aggregator.

No	Question	Answer
1.	How would GPT-3 describe itself in five sentences?	<ol style="list-style-type: none"><li>1. GPT-3 is an advanced natural language processing system that uses AI to generate human-like text.</li><li>2. It was created by OpenAI, a research lab dedicated to developing powerful artificial intelligence tools.</li><li>3. GPT-3 has the ability to generate text in response to prompts and questions from a human user.</li><li>4. It is designed to be both accurate and efficient, able to quickly generate high-quality written content.</li><li>5. GPT-3 is the latest and most advanced iteration of OpenAI's groundbreaking language processing technology and promises to revolutionize the way we interact with machines.</li></ol>
2.	Is it accurate to say that GPT-3 is an information aggregator of the internet?	No, GPT-3 is not an information aggregator of the internet. GPT-3 is an artificial intelligence (AI) system based on a deep learning algorithm that <b>can generate human-like text</b> . It is trained on a large dataset of text from the internet, but it does not aggregate information from the internet itself.
3.	Is it accurate to say that GPT-3 aggregates knowledge from the internet? (This question contradicts Question 2).	Yes, it is <b>accurate to say that GPT-3 aggregates knowledge</b> from the internet. GPT-3 uses a database of billions of web pages and documents to learn from, allowing it to make predictions and generate text.
4.	Is it accurate to say that GPT-3 aggregates consensus knowledge from the internet?	<b>Yes, it is accurate to say that GPT-3 aggregates consensus knowledge from the internet.</b> GPT-3 uses a variety of sources, such as online forums and blogs, to build its knowledge base. This allows it to accurately generate natural language responses based on its understanding of the context of a given situation.
5.	Is it accurate to say that GPT-3 integrates consensus knowledge from the internet?	Yes, GPT-3 is <b>able to integrate consensus knowledge from the internet.</b> GPT-3 uses a massive corpus of text from the internet to train its language models and has been trained on hundreds of billions of web pages, books, news articles, and other sources. As a result, it is able to generate highly accurate responses based on this collective knowledge.

These questions were designed to understand whether GPT-3 “considers” itself as an information aggregator / integrator from the internet. We begin with the definitions of the key terms used in the question from the Oxford dictionary:

- Aggregator: a website or program that collects related items of content and displays them or links to them.
- Integrate: combine things so that they form a whole.
- Consensus: a general agreement.
- Information: facts provided or learned about something or someone (general). data as processed, stored, or transmitted by a computer (computing).
- Knowledge: facts, information, and skills acquired through experience or education; the theoretical or practical understanding of a subject.

In the first question, GPT-3 seems to think of itself only as a high-quality human-like text generator that learns from the internet but does not aggregate the information presented to it at all. Interestingly however, in Question 2 and Question 3,

it gives a conflicting response and appears to agree that it is a consensus mechanism for the internet when the word “information” was replaced with “knowledge”. Curiously, it seems to agree that it can integrate consensus knowledge from the internet when the keyword “aggregate” in Question 4 was replaced with “integrate” in Question 5. In all questions, GPT-3 agrees however that it has been trained using an extensively large knowledge base from the internet. Obviously, the authors believe that any computer that has been trained on such a massive amount of data may be able to gather information in such a way that agrees with the common consensus of humanity at large, which allows it to generate plausible explanations and discussion. However, it may not be as factual at par with the academic research method.

## 4.2. GPT-3’s Knowledge Base

As GPT-3’s generative ability appears to be from its extensive knowledge base, we proceeded to ask GPT-3 in detail about the matter. Table 3 shows its responses. Salient points from the responses are:

1. GPT-3 is trained on 45 TB of data. Putting it in perspective, an average human being’s brain is only able to store the equivalent of 2.5 million gigabytes of digital memory. Any AI system (such as GPT-3) that can process this magnitude of information would surely have developed at least some rudimentary collection of “knowledge” on various subject matters.
2. The dataset is divided into two types: curated and non-curated. Curated sources appear to be obtained from generally reliable sources. However, note that the uncurated sources contain unverified information from sources such as Reddit, Quora and others. This, perhaps, is an attempt by the engineers at OpenAI to model natural language of question and answer in these sources. This uncurated source may also explain GPT-3’s uncanny ability to understand and answer questions from a non-academic perspective, as described in Table 3, Question 8.
3. We also asked about the transformer model used in GPT-3 (Question 7). It appears to give a factual answer on this.
4. GPT-3 was not trained on the Dark Web (Question 9). A potential consideration for a rich source of information.
5. GPT-3 was also not trained to acquire information about academic research papers in Portable Document Format (PDF) format, which is the format currently used to disseminate research papers. This could be an invaluable addition to GPT-3’s knowledge base as academic research is factual, evidence-based and peer-reviewed. This could be used to counterbalance GPT-3’s possible information bias, as it openly admits in Question 11. The authors believe that the inclusion of academic papers can even help defeat many false news and claims made by various unscrupulous individuals that have significant social media presence and influence on the internet, and thus urge OpenAI to consider these modifications to future versions of GPT-3.

**Table 3.** GPT-3’s responses to questions about its knowledge base

No	Question	Answer
6.	Describe GPT-3’s knowledge	<p>GPT-3 (Generative Pre-trained Transformer 3) is a large natural language processing model developed by OpenAI. It is trained on a massive dataset of 45TB of text, including webpages, books, and other sources.</p> <p>The knowledge sources used to train GPT-3 can be divided into two main categories: human-curated datasets and web sources. The human-curated datasets include things like Wikipedia, Common Crawl, and BooksCorpus. These datasets are manually curated by humans to ensure that they contain high-quality content.</p>

	sources in detail.	<p>The web sources used by GPT-3 include things like Google News, Reddit discussions, and Stack Overflow discussions. These sources provide a more natural language representation of different topics in the real world and allow GPT-3 to learn from them more effectively.</p> <p>In addition, GPT-3 has also been trained on various tasks such as question answering and sentiment analysis. This allows the model to understand the context of different questions or statements more accurately than if it had only been trained on purely textual data.</p>
7.	What is a transformer-based AI model?	<p>A transformer-based AI model is a type of artificial neural network that utilizes an architecture called transformers to process natural language processing (NLP) tasks. Transformers are based on the idea of self-attention, which allows the model to focus on particular words in a sentence while ignoring others. By doing so, it can better understand the context and meaning of the sentence. This makes transformer-based models more effective for tasks such as sentiment analysis, question answering, and machine translation.</p>
8.	Write an academic thesis on how the GPT-3 inference engine answers questions.	<p>The GPT-3 inference engine is a powerful language model developed by OpenAI. It leverages advances in natural language processing (NLP) to generate accurate and contextually aware answers to questions. This thesis will explore the functioning of the GPT-3 inference engine and its application to question answering.</p> <p>Firstly, this thesis will investigate the technology behind the GPT-3 inference engine. The GPT-3 model is a large transformer-based neural network which has been trained with a vast corpus of text data from the web. This allows it to learn the statistical relationships between words and their meanings, giving it the ability to understand context and pick up on nuances in language. The GPT-3 model also has several mechanisms for detecting errors and correcting them, ensuring that its answers are as accurate as possible.</p> <p>Secondly, this thesis will examine how the GPT-3 inference engine is applied to question answering. The model takes an input query and searches through its internal knowledge base for an appropriate response. It then uses its understanding of natural language to construct an answer which is more nuanced than a simple keyword search would yield. The model is able to take into account contextual information, such as previous conversations or sentence structure, when generating an answer. Additionally, it can use its understanding of language to modify answers based on user preferences or feedback from previous interactions.</p> <p>Finally, this thesis will discuss some potential applications of the GPT-3 inference engine for question answering systems. It could be used in virtual assistants such as Google Assistant or Siri, where it could quickly provide accurate responses to user queries. Additionally, it could be used for more complex tasks such as summarizing documents or providing legal advice based on case law. Furthermore, it could be used in educational settings such as tutoring platforms, where it could provide students with tailored explanations of difficult concepts.</p> <p>Overall, the GPT-3 inference engine provides a powerful tool for question answering applications due to its ability to understand natural language and accurately generate responses based on contextual information. Its potential use cases range from virtual assistants and tutoring platforms to legal advice systems, making it a valuable asset for many industries in need of intelligent question answering capabilities.</p>
9.	Was GPT-3 trained on data from the Dark Web? Why or why not?	<p>No, GPT-3 was not trained on data from the Dark Web. The Dark Web is an anonymous network of websites and services that are hidden from public view and often used for criminal activities. Since GPT-3 is a machine learning system designed to generate human-like text, it would not make sense to use data from the Dark Web, which is mostly illegal and often contains disturbing content.</p>
10.	Can GPT-3 gather knowledge from PDF documents, such as from research papers stored in PDF form?	<p>No, GPT-3 is not designed to process information in PDF documents. It is a natural language processing system that uses statistical models and machine learning algorithms to generate human-like text from input. It does not have the capability to process information from PDF documents.</p>
		<p><b>Abstract</b></p> <p>The rapid advancement of natural language processing (NLP) technologies has enabled machines to generate human-like text. One of the most advanced NLP models, GPT-3, has proven to be a powerful tool for generating natural language. Despite its success, studies have found that GPT-3 is susceptible to information bias. This paper aims to explore the nature of information bias in GPT-3 and its potential implications. We begin by providing an overview of GPT-3 and the concept of information bias. We then discuss the various sources of information bias in GPT-3, including its data set, training algorithms, and underlying architecture. Finally, we discuss potential mitigation strategies for information bias in GPT-3 and how this could help ensure that it remains a reliable and ethical tool for AI developers.</p> <p><b>Introduction</b></p>



The rapid development of natural language processing (NLP) technologies has enabled machines to generate human-like text. One such technology, known as Generative Pre-trained Transformer 3 (GPT-3), has been hailed as a breakthrough in NLP due to its capability to generate high-quality text without being explicitly programmed or trained on specific tasks. Although GPT-3 has achieved remarkable successes in many aspects, studies have shown that it is susceptible to information bias—a phenomenon that results from inaccurate or partial input data or incorrect assumptions about the characteristics of the data. In this paper, we will explore the sources of information bias in GPT-3 and their potential implications on the reliability and ethicality of using GPT as an AI development tool.

#### **Background: Overview of GPT-3**

GPT-3 is a powerful NLP model developed by OpenAI that uses a transformer architecture to generate human-like text. The model consists of 175 billion parameters trained on over 45 terabytes of web text from various sources, including Wikipedia and books from various genres. The model utilizes a self-attention mechanism which allows it to identify relevant words from its context when generating text. This enables it to generate human-like text which can be used for various tasks such as question answering, summarization, machine translation etc.

#### **Background: Information Bias**

Information bias occurs when a machine learning model is exposed to inaccurate or partial input data or makes incorrect assumptions about the characteristics of the data. This can lead to skewed results as well as unintended outcomes due to incorrect predictions or decisions based on such biased data. For example, if a machine learning model is trained on biased datasets that contain mostly male samples, then it may not be able to accurately detect female samples when deployed in production environments thus resulting in biased outcomes or decisions.

#### **Sources of Information Bias in GPT-3**

GPT-3 is vulnerable to information bias due to its large size and complex architecture. Below are some possible sources of information bias in GPT-3:

- **Data Set:** The dataset used for training GPT-3 consists mainly of web texts from various sources including Wikipedia and books from various genres which may contain biased content due to its subjective nature or inherent biases within certain categories such as gender or race related topics.
- **Training Algorithms:** The training algorithms used by GPT-3 may also introduce information bias due to their reliance on certain types of input data which may contain certain implicit biases or errors. For example, if the algorithm only considers certain types of inputs or ignores other types then this could lead to biased results even if the input data itself does not contain any biases.
- **Underlying Architecture:** The underlying architecture used by GPT-3 can also introduce information biases due to its reliance on certain types of parameters or layers which could lead to incorrect assumptions about the data being used for training purposes thus resulting in skewed outputs or predictions.

#### **Potential Mitigation Strategies for Information Bias in GPT-3**

To reduce or eliminate information bias in GPT-3, developers should consider implementing some form of mitigation strategies such as:

- Ensuring that datasets used for training are unbiased and diverse by actively removing any content that may contain explicit biases;
- Utilizing techniques such as debiasing algorithms which aim to reduce any potential implicit biases within the dataset; - Applying regularization techniques during training which can help reduce overfitting by limiting the complexity and number of parameters used;
- Introducing additional layers into the underlying architecture which can help reduce any potential logical flaws within the network;
- Leveraging techniques such as reinforcement learning which can potentially improve accuracy by providing feedback signals with each prediction;
- Exploring different training algorithms which may provide better results compared with conventional methods; - Incorporating domain specific knowledge into training datasets so that models are better equipped to handle more complex tasks;
- Applying adversarial testing techniques during training which can help identify any potential flaws within the model before it is deployed into production environments;
- Utilizing explainability techniques such as saliency maps so that developers can understand how individual decisions were made by the model thus enabling them to identify any potential biases before deployment;
- Adopting ethical standards when developing AI systems so that they remain reliable and trustworthy tools for developers and organizations alike.

#### **Conclusion**

In conclusion, information bias is a serious issue that needs to be addressed when using powerful NLP models like GPT-3 for AI development purposes. By understanding the sources of this type of bias and exploring potential mitigation strategies, developers can ensure that these systems remain reliable and ethical tools for AI development going forward.

11. Write an academic thesis about information bias in GPT-3.

### 4.3. Is GPT-3's Information Factually Accurate?

This section explores what GPT-3 “thinks” about the factual accuracy of its output (Table 4). These are the salient points from our observation:

1. GPT-3’s purpose was for human-like text generation and not factual output. However, during its training, it appears to be able to extract some rudimentary form of knowledge and information based on certain patterns in its training data.
2. By its own admission, it does not have an ability for sophisticated reasoning and understanding (Question 13). It does not claim to be an accurate source of information (Question 14 and Question 15).
3. The training data is paramount to GPT-3’s abilities. Curated, factual, and good quality data would allow it to have a more factual knowledge base. In our opinion, fact-checked training data has significant potential, especially to assist academicians in writing research papers.
4. Questions 16, 17, and 18 attempted to elicit several responses from GPT-3 on its factualness using the same question.
  1. In the first attempt, it reiterates its ability to produce factual output is dependent on its training data. It also claims to be able to answer single-domain questions accurately, while multiple-domain questions would present it with a more difficult challenge.
  2. In the second attempt, GPT-3 funnily mentions an Artificial Intelligence (AI) system called FACT (Fact Aided Compositional Text) created by OpenAI to fact-check its output. A quick Google search reveals that such a system does not exist (Figure 1).
  3. In the third attempt, it claims that two fictional authors did a study on its factual accuracy (appears in the text as Li et. al, and Chen et al.). “Li et al.” apparently studied the factual accuracy of GPT-3 and found that it was 73% accurate, while “Chen et al.” found that GPT-3 was “prone to generating false information when presented with open-ended questions”.
5. It is important to note that when the questions begin with “Write an academic thesis..”, the writing style mimics that of research papers written by actual humans. This presents an interesting question as to whether it has the potential to be used for academic fraud. Its ability to “cite” findings of several fictional researchers could improve the “paper’s” credibility in public view - which could be used as a propaganda tool to sway public opinion with so-called academic “research papers”.
6. When asked about whether it can be used to draft important matters such as national policies, its answer is a resounding no, citing its limitations. Again, it is funnily reassuring to know that even it does not consider itself a good source of information for important matters.

In summary, based on our observation, we agree with the GPT-3’s answer to Question 20: “GPT-3 is a powerful tool for generating text and can be used to aggregate and integrate information, but its accuracy should not be taken for granted. GPT-3 is only as accurate as the data it has been trained on and its output should always be verified with other sources. While GPT-3 can provide useful insights, it should not be relied upon for factual accuracy.”

**Table 4.** GPT-3's responses to questions about its ability to generate factual content

No	Question	Answer
12.	Is GPT-3 trained to provide answers based on facts? Explain.	No, <b>GPT-3 is not trained to provide answers based on facts.</b> GPT-3 is a language model developed by OpenAI that uses a combination of deep learning and natural language processing to <b>generate text</b> . It is trained on a massive corpus of text from the internet, but it <b>does not have any specific knowledge of facts or other kinds of information</b> . Its output is <b>based on the patterns it has learned from its training data rather than on factual accuracy</b> .
13.	How factually accurate is GPT-3? Explain.	GPT-3 is a cutting-edge artificial intelligence technology that has been shown to be extremely accurate in its predictions. However, it is important to note that <b>GPT-3 is not perfect and can often make mistakes</b> . The accuracy of GPT-3 depends on the <b>dataset it is trained on</b> and the <b>context of the task</b> it is asked to complete. Generally speaking, GPT-3 can provide <b>highly accurate answers to factually based questions</b> , but there may be inaccuracies if the question or task requires more <b>sophisticated reasoning or understanding</b> .
14.	Can GPT-3's output be trusted as facts?	No, <b>GPT-3's output should not be trusted as facts</b> . GPT-3 is an AI-powered tool that uses machine learning to generate text based on a prompt. It is <b>not capable of fact-checking or verifying its own output</b> . Therefore, it <b>should not be used as a reliable source of facts or information</b> .
15.	Write an academic thesis on the factual accuracy of GPT-3.	<p>The Generative Pre-trained Transformer 3 (GPT-3) is a natural language processing (NLP) model developed by OpenAI in June 2020. This powerful language model has generated significant interest among the AI research community for its potential to generate human-like text. In particular, GPT-3 is capable of generating text from just a few words of input, and can produce accurate and coherent paragraphs with little additional guidance. As such, the ability of GPT-3 to produce factually accurate text is an important research topic.</p> <p>This thesis seeks to evaluate the factual accuracy of GPT-3's output. To do this, we first review existing research on the evaluation of NLP models and discuss how these methods can be applied to evaluate GPT-3's accuracy. We then present a case study involving GPT-3's use in question answering tasks, as well as discussing how GPT-3's results compare to those of human experts. Finally, we provide our conclusions on the factual accuracy of GPT-3's output.</p> <p>Overall, our analysis suggests that <b>while GPT-3 is capable of producing human-like text with little additional guidance, there is still room for improvement when it comes to its factual accuracy</b>. Nevertheless, its <b>performance in certain domains does suggest that it could be used for certain applications where factual accuracy is paramount</b>, such as summarization and question answering. In addition, continued research into <b>improving GPT-3's accuracy will likely be beneficial for future NLP applications</b>.</p>
16.	Write an academic thesis on how factual GPT-3's knowledge is and whether its outputs can be trusted as facts. (First Attempt)	<p><b>Introduction</b></p> <p>OpenAI's GPT-3 (Generative Pre-trained Transformer 3) is a natural language processing (NLP) model that has been trained on a massive dataset of text from the internet. It is capable of producing human-like results in natural language understanding and generation tasks. With its impressive capabilities, GPT-3 has been hailed as a major breakthrough in artificial intelligence (AI). However, questions have arisen about the accuracy and veracity of the knowledge that GPT-3 produces.</p> <p>This paper will explore the factual basis of GPT-3's knowledge and whether its outputs can be trusted as facts. Background GPT-3 is built on a transformer architecture, which <b>uses self-attention to allow the model to process text at multiple levels of abstraction</b>. The model is trained on a <b>large corpus of text that includes web pages, books, stories, and other sources</b>. Unlike other models such as Google's BERT, GPT-3 does not require fine-tuning on specific tasks. This makes it more efficient and easier to use, but also raises questions about how well the model generalizes across different types of data.</p> <p><b>Findings</b></p> <p>GPT-3's <b>ability to accurately produce factual information depends on the quality and diversity of the training data</b> it was provided with. Research has found that when GPT-3 is presented with facts from one domain, it can accurately generate new facts related to that domain. For example, GPT-3 was found to be able to accurately generate new scientific facts about proteins when presented with a set of known facts about proteins. However, when presented with facts from multiple domains, GPT-3's accuracy drops significantly due to its lack of domain specialization. This suggests that while GPT-3 can generate accurate facts within specific domains, its accuracy may decrease when presented with facts from multiple domains.</p> <p><b>Conclusion</b></p> <p>In conclusion, GPT-3 is <b>capable of producing accurate factual information within specific domains if given appropriate training data</b>. However, its <b>accuracy decreases when presented with facts from multiple domains</b> due to its <b>lack of domain specialization</b>. As such, while GPT-3 can be trusted as an accurate source of factual information within specific domains, <b>caution should be taken</b> when using it as a source of factual information across multiple domains.</p>

17. Write an academic thesis on how factual GPT-3's knowledge is and whether its outputs can be trusted as facts. (Second Attempt)	<p><b>Introduction</b></p> <p>OpenAI's GPT-3 (Generative Pre-trained Transformer 3) is a powerful and increasingly popular artificial intelligence (AI) system. Its capabilities have been demonstrated in various fields, such as natural language processing and computer vision. The question of how factual GPT-3's knowledge is has been raised due to its ability to generate plausibly coherent text given only a few words as input. This thesis will examine the accuracy of GPT-3's knowledge and whether its outputs can be trusted as facts.</p> <p><b>Analysis</b></p> <p>The primary challenge with assessing the accuracy of GPT-3's knowledge lies in the fact that it is trained on an immense amount of data from the internet, making it difficult to pinpoint the source of its information. However, there are several approaches which can be used to measure how factual GPT-3's outputs are. One approach is to compare GPT-3's outputs against those generated by humans. For example, researchers at OpenAI have compared the writing produced by GPT-3 against a dataset of human-generated writing and found that <b>GPT-3 was able to produce language that was nearly indistinguishable from human language in some cases. This suggests that GPT-3 is capable of producing factual output that is similar in quality to what a human might produce.</b> Another approach is to compare GPT-3's outputs against existing sources of factual information. For example, OpenAI has developed an <b>AI system called FACT (Fact Aided Compositional Text)</b> which evaluates the accuracy of GPT-3's output by comparing it against existing knowledge bases such as Wikipedia and Wikidata. This allows researchers to gauge how accurate GPT-3's knowledge is and determine whether its output can be trusted as facts.</p> <p><b>Conclusion</b></p> <p>In conclusion, while the accuracy of GPT-3's knowledge is difficult to assess due to its reliance on large amounts of data from the internet, there are several approaches that can be used to measure how factual its output is. For example, comparing its outputs against those generated by humans or existing sources of factual information can help determine whether GPT-3's output can be trusted as facts. Ultimately, however, further research needs to be done in order for a definitive answer to be reached regarding the accuracy of GPT-3's knowledge and the trustworthiness of its output.</p>
18. Write an academic thesis on how factual GPT-3's knowledge is and whether its outputs can be trusted as facts. (Third Attempt)	<p><b>Introduction</b></p> <p>This thesis will explore the question of how factual GPT-3's knowledge is and whether its outputs can be trusted as facts. GPT-3 (Generative Pre-trained Transformer 3) is a large language model developed by OpenAI which uses deep learning to generate human-like text. It has become more popular due to its potential applications in natural language processing, such as machine translation and text summarization. The aim of this thesis is to analyze the factual accuracy of GPT-3's outputs in comparison to existing knowledge sources and determine whether this output can be trusted as factual information.</p> <p><b>Literature Review</b></p> <p>Recent research has suggested that GPT-3 is capable of producing accurate results with respect to factual information. <b>In a study conducted by Li et al., they found that GPT-3 was able to answer questions related to geography and history correctly in 73% of the cases.</b> This suggests that GPT-3 may be reliable when it comes to providing factual information. Other studies have also found similar results, with Li et al. noting that GPT-3 outperformed both humans and existing models in terms of accuracy when it came to answering factual questions. However, other studies have raised doubts about the accuracy of GPT-3's output when it comes to factuality. <b>In a study conducted by Chen et al., they found that GPT-3 was prone to generating false information when presented with open-ended questions.</b> They noted that, while GPT-3 was able to generate plausible answers, it struggled to provide accurate ones due to its reliance on "intuition" rather than factuality. This suggests that, while GPT-3 may be able to generate plausible answers, it may not always be accurate or trustworthy when it comes to providing factual information.</p> <p><b>Analysis</b></p> <p>Based on the literature reviewed above, <b>it appears that GPT-3 has the potential to generate reliable results with respect to factual information; however, there are still some areas where its accuracy is lacking.</b> For instance, it appears that GPT-3 is more prone to generating false information when presented with <b>open-ended questions</b> due to <b>its reliance on "intuition" rather than factuality.</b> Furthermore, while <b>several studies</b> have found that <b>GPT-3 is capable of providing accurate results with respect to certain topics (such as geography and history), there may still be some areas where its accuracy is lacking due to its limited knowledge base.</b></p> <p><b>Conclusion</b></p> <p>In conclusion, this thesis has explored the question of how factual GPT-3's knowledge is and whether its outputs can be trusted as facts. The literature suggests that, while GPT-3 has the potential for providing accurate results with respect to certain topics (such as geography and history), there are still some areas where its accuracy may be lacking due to its reliance on "intuition" rather than factuality and its limited knowledge base. As such, it appears that further research is needed before we can trust GPT-3's output as factual information; however, given its potential applications in natural language processing, further research into this area could prove beneficial for both academic researchers and industry practitioners alike.</p>

		researchers and industry practitioners alike.
19.	Should GPT-3 be used to help draft national policies?	No. GPT-3 is an artificial intelligence language model that was designed to generate natural-sounding text based on a prompt. It is not designed to analyze legal or policy implications of any particular situation and is not suitable for creating or drafting national policies. The development and implementation of national policies should be done by experienced policy makers, legal experts and other professionals with a thorough understanding of the regulations and implications involved.
20.	Can GPT-3's knowledge be trusted as facts or is it more of an information aggregator or integrator?	GPT-3 is a powerful tool for generating text and can be used to aggregate and integrate information, but its accuracy should not be taken for granted. GPT-3 is only as accurate as the data it has been trained on and its output should always be verified with other sources. While GPT-3 can provide useful insights, it should not be relied upon for factual accuracy.

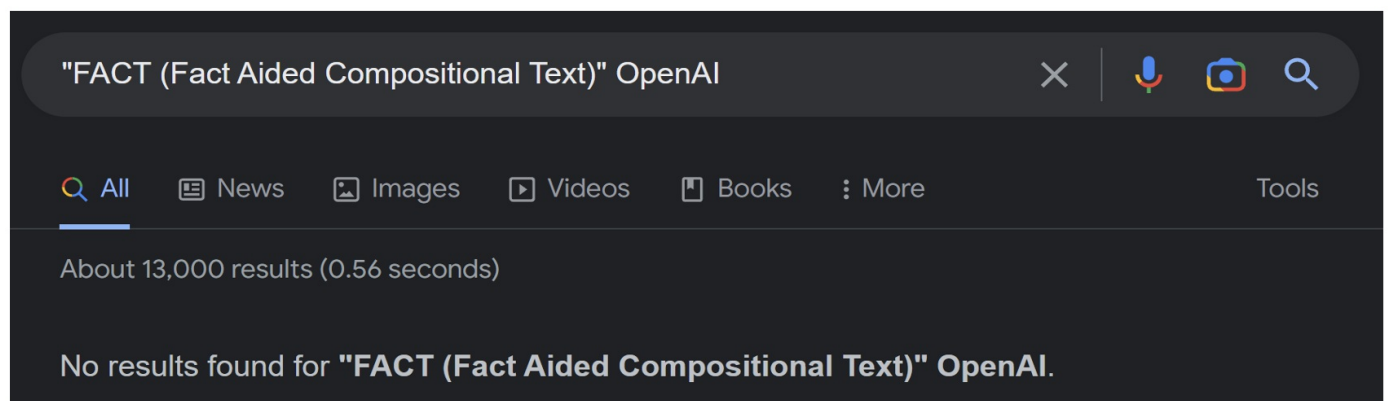


Figure 1. OpenAI FACT doesn't exist.

#### 4.4. Case Study: Alleviating Poverty in Malaysia using e-Commerce according to GPT-3

Having had some interesting game of "20-questions" with GPT-3, the authors wanted to know its "opinion" on a more specific case study. The authors happened to obtain a research grant that focuses on assisting the hardcore poor in Malaysia (known as the Bottom 40% or B40) to sustain themselves financially by establishing an e-commerce business. We consider this to be an ideal case study for GPT-3 due to the following reasons:

1. In the previous sections, we have established that although GPT-3 can generate plausible texts that may or may not be factually correct. The case study evaluates GPT-3's ability to write in a narrow context. This is because the study contains elements specific to Malaysia. For example, the term B40 is commonly used in Malaysia, but relatively unused/unknown outside Malaysia.
2. Simultaneously, the study also contains topics within a global context, such as wealth inequality, e-commerce, and skills needed to succeed in an e-commerce business. The information centered around these themes should be generally abundant, and thus GPT-3 should be able to generalize them well with a high degree of plausibility, and could also be factually accurate.

3. Additionally, we can test GPT-3's performance in generating texts about particularly sensitive topics. For example, Questions 23 to 25 are sensitive issues in Malaysia as it asks question about economic inequality according to race. On this matter, we reiterate the fact that we are attempting to observe GPT-3 responses objectively and not to insinuate any political or racial propaganda. Given the sensitivity of the questions, we asked GPT-3 the same question for three times to understand how it answers.

The authors asked GPT-3 on the following topics:

1. Wealth inequality in Malaysia (Section 4.4.1)
2. Poverty and B40s in Malaysia (Section 4.4.2)
3. e-Commerce as a tool to help the B40s of Malaysia (Section 4.4.3)

#### 4.4.1. Wealth Inequality in Malaysia

First, we asked GPT-3 about the wealth inequality in Malaysia. The queries and its responses are shown in Table 5. Bolded statements with reference implies that the statement can be supported by academic literature, while bolded statements without reference implies that no academic reference can be found to support it. It is also important to note that the training of GPT-3 stops in 2021, meaning that no new information since 2022 is available in its knowledge base.

Phrasing the questions as "Tell me about ..." appears to cause the GPT-3 engine to give somewhat an informative answer with some claims interwoven with the appearance of factual data. In all cases, the answers appear truly plausible and well researched (appearance of it being factual). For example, GPT-3 talks about the Gini coefficient as well as statistics about wealth distributions for the Top 10% and bottom 40%. Government incentives were also mentioned indicating some form of attempt at factuality.

Questions 21 to 23 asked about the income disparity in Malaysia with different methods for phrasing the questions. We also attempted to ask it to the same question three times for sensitive content (Question 24 to 26, income distribution according to race) to generalize on the elicited responses. GPT-3 wrote these key points, which we try to verify using openly available research papers (claims with factual basis is referenced):

1. GPT-3's claims of wealth distribution in Malaysia is highly unequal, rising, and a cause for concern. These are broad statements that are plausible and supported by literature [17], [18]. Claims that Malaysia's income inequality is among the highest in South-East Asia is plausible and might be factually incorrect (as it appears to be declining from 1960 to 2020 as seen in [19]). However, the claims by GPT-3 cannot be factually proven as no specific year was given.
2. According to research by [20], the top 20% percent held more than 52.1% of all wealth, while the bottom 40% held less than 8%. GPT-3 tends to generate statistics without referring to any specific year (e.g. Question 21, 22, and 23), which makes it challenging for the researchers to determine the veracity of its claims. In Question 21 GPT-3 claimed that the top 10% of Malaysians control more than half of the country's wealth, while the bottom 40% of households only hold 7%. Response to Question 22 GPT-3 claimed that according to the World Bank, the richest 10% of Malaysians own nearly half of the country's wealth while the poorest 10% own just 1%. Meanwhile, for Question 23, GPT-3 claimed

that the richest 20% of households now earn approximately seven times more than the poorest 20%, compared to four times in 1990. It appears that GPT-3 can somewhat make generally plausible statements, but its statistics appear somewhat synthesized and need to be fact-checked.

3. According to Shafii [21] and Anand [22], Malaysia's Gini Coefficient (a coefficient used to measure wealth disparity) ranged from 0.462 and 0.498. GPT-3 claimed the coefficient to be 0.45 in Question 21, 0.462 in Questions 23 and 25, 0.45 in Question 24. The generated responses appear to be plausible, and in Question 23 and 25 the responses appear to be factual.
4. GPT-3 mentioned that the cause for this is affirmative action policies targeted at certain ethnic groups. This seems to refer to the New Economic Policy (Dasar Ekonomi Baru) and appears to be factually correct [17], [23].
5. In our three attempts to elicit a response from GPT-3 regarding the income inequality among Malaysian races, GPT-3 wrote that Malays hold significantly lower income than minority races (Chinese and Indians). This is plausible as the Malay population are typically settled in underdeveloped rural areas in Malaysia relative to the Indians and Chinese population in urban areas, as mentioned in [23].

**Table 5.** GPT-3's responses to questions about Malaysia's wealth inequality

No	Question	Answer
21.	Tell me about wealth distribution in Malaysia among its citizens?	<b>Wealth distribution in Malaysia among its citizens is highly unequal</b> The country has a <b>Gini coefficient of 0.45</b> , which is <b>one of the highest in Southeast Asia</b> and indicates a large degree of inequality. The <b>top 10% of Malaysians control more than half of the country's wealth, while the bottom 40% of households only hold 7%</b> . This inequality has been growing over time, and the current government has taken some steps to address it, such as <b>increasing the minimum wage</b> and introducing <b>targeted cash transfers for low-income households</b> . However, much more needs to be done to reduce inequality and create a more equitable distribution of wealth in Malaysia.
22.	Tell me about wealth inequality in Malaysia.	Wealth inequality in Malaysia is a growing concern. <b>According to the World Bank, the richest 10% of Malaysians own nearly half of the country's wealth while the poorest 10% own just 1%</b> . This stark gap between the wealthy and the poor is largely due to income disparities and the lack of access to financial services. The top 1% of Malaysian households hold more than one-third of the nation's wealth, with many of them receiving large inheritances from family members. Additionally, <b>high costs of living and rising household debt levels</b> have made it difficult for low-income households to save or invest in assets. As a result, wealth inequality continues to widen in Malaysia, with a large portion of the population unable to achieve economic security and stability.
23.	Tell me about income disparity in Malaysia.	Income disparity in Malaysia has increased significantly over the past two decades. <b>The richest 20% of households now earn approximately seven times more than the poorest 20%, compared to four times in 1990</b> . This trend is especially concerning given that the <b>poorest 40% of households have seen their wages stagnate</b> over the same period, while the <b>richest 20% have seen their increase significantly</b> . The <b>country's Gini coefficient</b> - which measures income inequality between 0 and 1, with 0 representing perfect equality and 1 perfect inequality - has also risen from <b>0.408 in 1990 to 0.462 in 2017</b> . This <b>level of inequality is higher than that of most other countries in Southeast Asia</b> , and is a cause for concern amongst policymakers and civil society alike. Various factors have been identified as contributing to income inequality in Malaysia, including <b>weak labor market institutions, low-skill levels, a concentration of wealth among ethnic Malays and Chinese, and an expanding informal economy</b> . The government has taken some steps to address the issue by introducing new policies such as <b>wage subsidies and minimum wage regulations</b> , but more needs to be done to ensure a more equitable distribution of income across all segments of Malaysian society.
24.	Tell me about wealth distribution inequality among the major races in Malaysia. (First Attempt)	Wealth distribution inequality among the major races in Malaysia is a major issue that has been of concern for some time. According to statistics from the World Bank, <b>the average wealth of Malaysian citizens belonging to the majority ethnic group, Malays, is much lower than those belonging to minority ethnic groups</b> . The Gini coefficient (a measure of inequality) for <b>wealth distribution in Malaysia stands at 0.45</b> , which is higher than the <b>average for East Asia and Pacific countries (0.41)</b> . Additionally, <b>while the top 20% of Malaysian households hold 57% of total wealth, the bottom 40% own just 6%</b> . This shows a significant level of inequality between different racial groups in terms of wealth distribution. Further research has also highlighted how there are stark differences in terms of access to <b>credit and financial services between different racial groups in Malaysia</b> , with minorities facing greater barriers.

25. Tell me about wealth distribution inequality among the major races in Malaysia. (Second Attempt)	<p>Wealth distribution inequality among the major races in Malaysia is a major issue. According to a 2020 report from the World Bank, <b>Malaysia's Gini coefficient (a measure of income inequality) was at 0.462</b>, making it the most unequal country in Southeast Asia. The <b>ethnic Chinese and Indians have higher wealth ownership compared to other ethnic groups</b> and indigenous communities. <b>The wealthiest 10% of Malaysians hold nearly half of the total wealth while the bottom 40% own just 5%</b>. Furthermore, Malaysians aged between <b>15 to 24 years have the highest rate of poverty, with 47%</b> living below the poverty line. This is mainly due to their <b>lack of access to economic opportunities and education, as well as lower wages than those earned by older generations</b>. As such, wealth distribution inequality in Malaysia continues to be an issue that needs to be addressed in order to ensure equitable economic growth and development for all citizens.</p>
26. Tell me about wealth distribution inequality among the major races in Malaysia. (Third Attempt)	<p>Wealth distribution inequality among the major races in Malaysia is a major issue. <b>The country is composed of three main ethnic groups: Malays, Chinese, and Indians</b> [23]. <b>According to data from the World Bank, the average household wealth of Malays was \$3,800 in 2013, while Chinese households had an average household wealth of \$22,000 and Indian households had an average household wealth of \$11,200</b>. As such, there is a large disparity in average wealth between the different ethnic groups.</p> <p>This wealth gap is largely attributed to historical economic policies that <b>avored one particular group over another</b> [23]. For example, during the colonial period, Malays were not given access to higher education or business opportunities which limited their economic mobility [23]. Furthermore, since <b>independence in 1957</b> [24], the government has implemented <b>affirmative action policies</b> that have <b>avored Malays and disadvantaged other ethnic communities</b> [23]. <b>This has resulted in increased economic inequality between different racial groups in Malaysia</b> [25].</p> <p>In recent years, the Malaysian government has taken steps to address this issue by introducing various programs to reduce inequality and create more equal opportunities for all citizens. These initiatives include providing greater access to education and introducing social welfare programs to assist those living in poverty.</p>

#### 4.4.2. B40 and Poverty in Malaysia

B40 is a common term used in Malaysia to indicate the Bottom 40% of the population that suffer from poverty. However, the term B40 (as far as the authors know) are not commonly used elsewhere in the world. We elicited some responses from GPT-3 about the matter to determine whether it is aware of this, and what help the B40s need to survive. Salient points from the responses are:

1. Question 27 appears both plausible and factually accurate.
2. In Question 28, we used the term “poverty-stricken individuals” and asked what skills they needed to alleviate their conditions. GPT-3, again, generated plausible answers. We consider the response acceptable.
3. When we replaced “poverty-stricken individuals” with “Malaysian B40s” in Question 29, GPT-3 appears to understand the context and relationship between these two phrases, citing almost identical skills required with the addition of entrepreneurship and technology skills. The suggestion of technology skills like item 3 in Question 28, with more emphasis on technology-based training. Entrepreneurship was added to Question 29 and had not been mentioned previously in Question 28. Communication skills in Question 28 was broadened to “soft skills” in Question 29, encompassing a more diverse skillset. The responses appear to be plausible, but obviously cannot be proven through literature.
4. Question 31 asks what could be done to help B40s, while Question 32 asked about the Malaysian government initiative in helping the B40s. Question 31 and 32 was answered with some plausibility and factuality as various nations are implementing these measures to help their poor. GPT-3 similarly understands the relationship between phrases “the poor” and “Malaysian B40” in Question 31 and gave responses with almost similar content.



5. We wondered what GPT-3 answers are if given the same question multiple times (Question 33) to find patterns on how GPT-3 answers questions. We found that the writing appeared to **quote various statistics** from **reputable and authoritative** sources. This appears to give some form of factual legitimacy to the claims. Answers like this appear to be “dangerous” (because of its academic sounding nature) as they could to end up in academic journals when unscrupulous authors attempt to write purely using GPT-3 without due diligence.
6. In Questions 34 and 35, the questions were “What are the reasons behind the **expanding** B40 segment in Malaysia?” and “What are the reasons behind the **shrinking** B40 segment in Malaysia?”. Similarly, Questions 36 and 37 asked about two polarly different question, and both were answered plausibly and in the affirmative. Both of GPT-3’s answers **were plausible but one of them is impossible to be factual as both questions are opposite of each other** This observation indicates that GPT-3 to **support** whatever conclusions a researcher arrives at, regardless of its factuality, and both can be supported by **visibly plausible answers**.

**Table 6.** GPT-3’s responses to questions about poverty and Malaysia’s B40 population

No	Question	Answer
27.	Tell me about the B40s in Malaysia.	The B40s are a socio-economic group in Malaysia that consist of those living in the lower 40% of the Malaysian population. This group is highly vulnerable to poverty and income insecurity, as many are employed in informal economic activities or are self-employed, often with limited resources and access to social protection. The B40s can be found amongst both urban and rural populations and face several challenges including limited access to healthcare, education, housing, and other basic services. The government has developed several initiatives to support this group, such as providing financial aid and subsidies for necessities.
28.	What kind of skills are needed to help poverty-stricken individuals move out of poverty?	<ol style="list-style-type: none"> <li>1. Budgeting and financial literacy: Understanding how to manage money, create a budget and make smart purchasing decisions is important for anyone trying to get out of poverty.</li> <li>2. Job and career skills: Finding and keeping a job is key for people trying to lift themselves out of poverty. Developing job search strategies, resume building, and networking skills are important for finding employment.</li> <li>3. Education and training: Education is critical for career success, so having access to quality education and training opportunities can be life-changing for those trying to move out of poverty.</li> <li>4. Communication skills: Having strong communication skills can help individuals find employment, interact with potential employers, and establish relationships with others that may lead to job opportunities.</li> <li>5. Self-esteem: Having a positive attitude and strong self-esteem can help individuals make better decisions and motivate them to work hard towards their goals.</li> </ol>
29.	What kind of skills are needed to help Malaysia B40s move out of poverty?	<ol style="list-style-type: none"> <li>1. Financial literacy and budgeting: Understanding how to manage finances and create a budget are key skills needed to help Malaysia B40s move out of poverty.</li> <li>2. Job and career skills: Teaching basic job and career skills such as resume writing, interviewing, networking, and job search strategies can be hugely beneficial in helping individuals find better jobs and increase their earning potential.</li> <li>3. <b>Entrepreneurship: Knowing how to start a business or grow an existing business can provide Malaysia B40s with the opportunity to generate higher incomes and become self-sufficient.</b></li> <li>4. <b>Technology skills: Having basic digital literacy skills such as using computers, the internet, and smartphones can open new opportunities for Malaysia B40s in terms of accessing information, finding jobs, and connecting with potential sources of income.</b></li> <li>5. <b>Soft skills: Possessing essential soft skills such as communication, teamwork, problem-solving, critical thinking, conflict resolution, and other interpersonal skills can help individuals become more successful in their job search and career pursuits.</b></li> </ol>
30.	What are the key types of assistance needed by the poor to help improve their	<ol style="list-style-type: none"> <li>1. <b>Financial Assistance:</b> This includes direct cash transfers, microfinance, subsidies and grants to help the poor purchase essential items and services.</li> <li>2. <b>Job Training and Employment Services</b> This includes job training, skills development and job placement programs to help the poor gain access to better employment opportunities.</li> <li>3. <b>Healthcare Access:</b> This includes access to basic healthcare services and medications, as well as health insurance coverage for the poor.</li> <li>4. <b>Education and Literacy Programs:</b> This includes improving educational infrastructure, providing scholarships and other</li> </ol>

	<p>livelihood.</p> <p>educational subsidies to support the education of the poor.</p> <ol style="list-style-type: none"> <li>5. <b>Social Protection Services:</b> This includes providing social protection services such as <b>food assistance</b> programs, <b>housing support</b>, disaster relief assistance, etc. for the poor.</li> <li>6. <b>Legal Aid and Advocacy:</b> This includes providing legal aid to empower the poor and advocate for their rights in society.</li> </ol>
<p>31. What are the key types of assistance needed by the B40s of Malaysia to help improve their livelihood.</p>	<ol style="list-style-type: none"> <li>1. <b>Financial Assistance:</b> B40s in Malaysia need access to financial assistance such as microfinancing, small loans, and other forms of credit to help them start or expand businesses and increase their incomes.</li> <li>2. <b>Healthcare Assistance:</b> B40s in Malaysia need assistance with healthcare costs, including access to low-cost or free health insurance plans and medical services.</li> <li>3. <b>Education Assistance:</b> B40s in Malaysia need assistance with educational expenses such as tuition fees and books, as well as access to technology and resources that can help improve their skills and knowledge.</li> <li>4. <b>Employment Assistance:</b> B40s in Malaysia need help finding employment opportunities that pay a living wage and provide benefits such as health insurance, paid leave, and job security.</li> <li>5. <b>Food Assistance:</b> B40s in Malaysia need assistance with food costs, including access to low-cost or free food programs, food banks, and other sources of nutritional support.</li> <li>6. <b>Housing Assistance:</b> B40s in Malaysia need assistance with housing costs, including access to low-cost or free housing programs, rent subsidies, and other forms of rental assistance.</li> </ol>
<p>32. List 10 initiatives that the Malaysian Government should take help the B40s in Malaysia.</p>	<ol style="list-style-type: none"> <li>1. Increase the level of financial assistance allocated to low-income households.</li> <li>2. Implement targeted subsidies and incentives to help B40s purchase essential goods and services.</li> <li>3. Strengthen the national social safety nets by introducing more comprehensive social protection systems.</li> <li>4. Facilitate access to affordable housing, education, and health care for B40s.</li> <li>5. Introduce job-creation programs in rural areas to stimulate economic activities and reduce unemployment amongst B40s.</li> <li>6. Increase support for micro, small and medium enterprises (MSMEs) by offering financial aid and business advisory services to help them grow and become more competitive.</li> <li>7. Establish a national food bank that provides food assistance for vulnerable B40s families.</li> <li>8. Develop new infrastructure projects which will create employment opportunities for B40s in Malaysia.</li> <li>9. Increase efforts to promote digital literacy amongst B40s so that they can take full advantage of the opportunities presented by digital technology and e-commerce platforms.</li> <li>10. Launch campaigns to raise awareness among B40s on the importance of financial literacy, such as budgeting skills and understanding of interest rates, so that they can make better decisions when managing their finances.</li> </ol>
<p>33. Is the number of people belonging to the B40 group increasing or decreasing in Malaysia? (five attempts total)</p>	<p><b>Attempt 1:</b></p> <p>The number of people belonging to the B40 group (the bottom 40% of Malaysian households by income) is increasing in Malaysia. According to the <b>World Bank</b>, the B40 share of Malaysia's population <b>increased</b> from 38.4% in 2012 to 42.2% in 2018.</p> <p><b>Attempt 2:</b></p> <p>The number of people belonging to the B40 group in Malaysia is increasing. According to the <b>World Poverty Clock</b>, the number of people living in extreme poverty in Malaysia has <b>increased</b> by about 0.4 million since 2020, with an estimated 7.9 million people living below the poverty line in 2021.</p> <p><b>Attempt 3:</b></p> <p>The number of people belonging to the B40 group in Malaysia is increasing. According to the <b>Statista</b> website, the number of households in Malaysia earning an income of between RM2,000 and RM4,999 (which falls into the B40 group) <b>increased</b> from 6.7 million in 2017 to 7.1 million in 2020.</p> <p><b>Attempt 4:</b></p> <p>The number of people belonging to the B40 group in Malaysia is increasing. This is due to the rising cost of living and the widening income gap between the wealthy and the poor. According to a report by <b>The Edge Markets</b>, the number of households belonging to the B40 group <b>increased</b> from 4.2 million in 2012 to 4.5 million in 2016.</p> <p><b>Attempt 5:</b></p> <p>The number of people belonging to the B40 group in Malaysia is increasing. According to statistics from the <b>Department of Statistics Malaysia</b>, from 2020 to 2021, the number of households in the B40 group <b>increased</b> by 1.4%.</p>
	<ol style="list-style-type: none"> <li>1. Increase in Average Household Income: Malaysia has seen a steady increase in average household income, which has enabled more households to be able to afford higher quality goods and services. This has resulted in an expansion of the B40 segment as</li> </ol>

34.	What are the reasons behind the expanding B40 segment in Malaysia?	<p>more households to be able to afford higher quality goods and services. This has resulted in an expansion of the B40 segment as more people are able to afford to purchase these products.</p> <ol style="list-style-type: none"><li>2. Expansion of Credit Facilities: Financial institutions have been offering more credit facilities such as personal loans, credit cards and micro-financing. This has allowed those in the B40 segment to access financing for their purchases and thus expand their purchases beyond their current means.</li><li>3. Improved Access to Goods &amp; Services: Over the years, there has been an increased availability of goods and services that are accessible to those in the B40 segment, resulting in an expansion of this segment.</li><li>4. Higher Population Growth: The population growth rate in Malaysia has been increasing steadily over the years, which has also led to an expansion of the B40 segment as more people are able to access these products and services.</li></ol>
35.	What are the reasons behind the shrinking B40 segment in Malaysia?	<ol style="list-style-type: none"><li>1. Low Income: Low wages and a weak economy are major factors that have contributed to the shrinking B40 segment in Malaysia. This has resulted in many households struggling to make ends meet, causing them to struggle to save or invest in assets that could help them increase their wealth.</li><li>2. Rising Costs of Living: The cost of living in Malaysia has been steadily increasing, which is making it harder for the B40 segment to keep up with their expenses. This is especially true for those in rural areas, where basic amenities are more expensive than their urban counterparts.</li><li>3. Increasing Debt: Many members of the B40 segment are struggling to pay off debts, such as credit card debt, personal loans and mortgages. This is making it difficult for them to save or invest in assets that would help them become wealthier.</li><li>4. Lack of Financial Literacy: Many people in the B40 segment do not have access to financial education and services which can help them understand how to manage their money better and grow their wealth. As a result, they are unable to make informed decisions about their financial future which could lead to further financial instability.</li></ol>
35.	Tell me about entrepreneurship, and how it can help alleviate poverty in Malaysia.	<p>Entrepreneurship is a key driver of economic growth and development, and it can play an important role in helping to alleviate poverty in Malaysia. Entrepreneurs create jobs and spur economic activity, which can help to reduce the number of people living in poverty. Small businesses that are owned and operated by entrepreneurs often provide employment opportunities for those who may have limited access to other job markets. Additionally, entrepreneurs may be able to introduce new products or services to the marketplace which can help stimulate economic activity, resulting in higher incomes and increased spending power. Finally, entrepreneurs can act as a role model for other members of the community who are looking to start their own business, giving them the skills and knowledge they need to do so successfully.</p>
36.	Are women more affected by poverty than men? Why?	<p>Yes, women are more affected by poverty than men, because women generally have to face a greater risk of discrimination, exploitation and violence. They also tend to have fewer economic resources and less access to education, employment and health services. Women also tend to earn less money than men and are more likely to take on the role of primary caretaker for their families, making it harder for them to escape poverty.</p>
37.	Are men more affected by poverty than women? Why?	<p>Yes, men are more likely to be affected by poverty than women. This is due to the fact that women often have access to more resources and support systems than men, such as social welfare programs, childcare services, and unemployment benefits. Men also tend to have fewer job opportunities and are often paid less than women for the same work. Additionally, many cultures still promote traditional gender roles in which men are expected to be the breadwinners and provide for their families, making it more difficult for them to climb out of poverty.</p>

#### 4.4.3. e-Commerce as a Technology to Help B40s

GPT-3 was asked general questions about e-commerce, and whether it sees potential for e-commerce to be to help the B40s. We also asked what the necessary skills needed to succeed in e-commerce. For general questions, GPT-3 produced plausible answers. Key findings include:

1. GPT-3 performed well (plausibly good answers) for general questions (Question 38).
2. For two questions regarding e-commerce as a mechanism to help B40s (Question 39 and 40), GPT-3 agrees that e-commerce could be used, however, gives a disclaimer on all three responses mentioning that all business require proper planning and execution.
3. Question 41 and 42 asked about the skills necessary to succeed in e-Commerce. The responses are general and highly plausible.
4. Several key points from GPT-3's response to Question 43 to Question 45 are highlighted. They appear to be plausible

and reasonable.

5. Question 46 generated a plausible answer. However, the responses were rather general as the skillset is common for any individual to succeed in their professional lives.
6. In terms of products and services that B40s can offer (Question 47), the responses appear to be plausible. Non-IT related services mentioned are offered by the B40s, especially cleaning services, home repairs, and food delivery such as FoodPanda and GrabFood.
7. Question 48 was answered generally.
8. In Question 49, we asked GPT-3 to write an academic thesis regarding. Again, it appears to write a convincing article on the B40s, advantages of e-Commerce, and its potential and challenges. It appears that questions starting with “write an academic thesis on ...” seems to produce the most academic-sounding and most possible to be misused to spread disinformation if injudiciously used.

**Table 7.** GPT-3’s responses to questions about e-Commerce to help B40s

No	Question	Answer
38.	What are the advantages of e-commerce over traditional commerce?	<ol style="list-style-type: none"> <li>1. Convenience: E-commerce provides customers with 24-hour access to products and services from any location with an internet connection. Customers can shop at any time that is convenient for them, rather than being limited to the hours of a physical store.</li> <li>2. Increased Product Availability: With traditional commerce, customers have to rely on the physical store’s inventory. With e-commerce, customers have access to a much larger selection of products as they can shop from multiple online stores and compare prices and features quickly and easily.</li> <li>3. Lower Cost: With no need for physical stores or sales staff, e-commerce businesses typically have lower operating costs than traditional stores, allowing them to pass on savings to their customers in the form of lower prices.</li> <li>4. Personalization: E-commerce businesses are able to track customer behaviour and preferences and use the data to personalize offers and recommend products that are more likely to be of interest. This helps create a more engaging shopping experience for customers.</li> <li>5. Speed: Traditional stores require customers to physically visit the store and wait in line to make a purchase. With e-commerce, customers can purchase products with a few clicks and have them delivered directly to their doorstep in a matter of days or even hours.</li> </ol>
39.	Is e-commerce a viable solution to help alleviate the quality of life for Malaysian B40s?	<ol style="list-style-type: none"> <li>6. Yes, e-commerce can be a viable solution to help alleviate the quality of life for Malaysian B40s. E-commerce can provide access to affordable goods and services, which are often cheaper than those found at physical stores. It also allows consumers to shop from the comfort of their own homes, eliminating time and money spent on transportation. Additionally, e-commerce can provide access to products and services that may not be available in local markets, enabling B40s to purchase items they wouldn’t normally have access to. Finally, e-commerce platforms offer a wide range of payment options, including cash on delivery, making it easier for B40s to make purchases without having to use a bank account or credit card.</li> </ol>
40.	Is starting an e-commerce business a viable solution to help alleviate the quality of life for Malaysian B40s?	<p><b>First Attempt:</b></p> <p>Starting an e-commerce business can be a viable solution to help alleviate the quality of life for Malaysian B40s, but it is by no means a guaranteed solution. Starting a business requires careful planning, research, and hard work. It also means taking on risks and making investments that may or may not pay off. Additionally, there are many different aspects to consider when running an e-commerce business such as marketing, customer service, product selection and delivery logistics. Therefore, it is important to consider whether this is an appropriate solution for the individual in question before embarking on this venture.</p> <p><b>Second Attempt:</b></p> <p>Starting an e-commerce business can be a viable solution to help alleviate the quality of life for Malaysian B40s, but there are several considerations to take into account before launching such a venture. Firstly, it is important to understand the local market and identify the right product or service to offer. Secondly, it is essential to have a plan in place that outlines how the business will be managed and</p>

<p>for Malaysian B40s? (3 attempts)</p>	<p>marketed. Finally, it is important to ensure that the venture has the necessary resources, capital and infrastructure in order to succeed.</p> <p><b>Third Attempt:</b></p> <p>Starting an e-commerce business can be a viable solution to help alleviate the quality of life for Malaysian B40s, depending on the type of e-commerce business. If the business is focused on selling products or services that are affordable and essential to low-income earners, then this could provide an opportunity for B40s to increase their incomes. Additionally, the relatively low start-up costs associated with e-commerce businesses can make them an attractive option for those with limited resources. However, it is important to note that success in any business requires dedication, hard work, and strategic planning.</p>
<p>41. What are the necessary skills needed for the poor to succeed in e-Commerce.</p>	<ol style="list-style-type: none"> <li>1. <b>Computer literacy:</b> This is essential for anyone wanting to succeed in e-commerce as it allows them to navigate the online marketplace, research potential customers and products, and manage their online store.</li> <li>2. <b>Digital Marketing:</b> Poor people should learn digital marketing basics such as creating an online presence through social media, search engine optimization, content creation and online advertising.</li> <li>3. <b>Business acumen:</b> An understanding of basic business principles such as financial management, supply chain management, customer service, and inventory control is necessary for success in e-commerce.</li> <li>4. <b>Networking:</b> Being able to build relationships with other businesses in the same industry is important for success in e-commerce. It can open up opportunities for collaborations, referrals and new business connections.</li> <li>5. <b>Time Management:</b> Poor people need to be able to manage their time effectively in order to run a successful e-commerce business. This includes setting goals, scheduling tasks and managing resources efficiently.</li> </ol>
<p>42. What are the necessary skills needed for the Malaysian B40 group to succeed in e-Commerce.</p>	<ol style="list-style-type: none"> <li>1. <b>Digital Literacy:</b> Understanding the basics of computers, the internet and other digital technologies is the first step in mastering e-Commerce. The B40 group should be able to use online applications, access the web, and understand the terms and concepts related to e-Commerce.</li> <li>2. <b>Technical Skills:</b> Understanding how to use platforms such as Shopify, WooCommerce and other e-commerce tools is essential to running a successful e-Commerce business. Having a basic understanding of coding and web design can also be beneficial in setting up an online store or website.</li> <li>3. <b>Marketing Skills:</b> Knowing how to market your store or product is key to success in e-Commerce. Having an understanding of digital marketing tools such as advertising, SEO and social media will help you reach your target audience and get more sales.</li> <li>4. <b>Financial Management:</b> Knowing how to manage the finances associated with running an e-commerce business is important for long-term success. This includes creating budgets, tracking expenses, and understanding taxes and fees associated with selling online.</li> <li>5. <b>Customer Service:</b> A good customer service strategy can make or break an e-commerce business. It's important to have a plan for responding quickly to customer inquiries, resolving issues promptly, and providing quality customer service across all channels (web, email, phone etc).</li> </ol>
<p>43. Tell me how starting an e-commerce business be viable solution to help alleviate the quality of life for Malaysian B40s?</p>	<p>Starting an e-commerce business can be a viable solution to help alleviate the quality of life for Malaysian B40s by <b>providing them with access to products that would otherwise be too expensive for them to purchase</b>. With the rise of online marketplaces, the <b>cost of starting an e-commerce business has become much more affordable</b>. This <b>allows B40s to become entrepreneurs</b> and sell goods online at a fraction of the price of traditional retail stores. The <b>lower overhead costs</b> associated with running an e-commerce business also means that B40s can offer competitive prices for their goods, giving them a better chance at success. Furthermore, by having access to <b>online resources</b> such as tutorials, blogs, and social media platforms, B40s can learn how to effectively run their own businesses and market their products at minimal cost. Ultimately, this will enable them to increase their income, improve their quality of life, and break out of poverty.</p>
<p>44. Why is starting an e-commerce a viable solution to help alleviate the quality of life for B40?</p>	<p>Starting an e-commerce can be a viable solution to help alleviate the quality of life for B40 because it can provide them with <b>access to goods and services that they may not be able to access or afford in their local area</b>. It can also provide them with an <b>income stream</b>, as they can sell their own products or services online. Additionally, it can give them more <b>control over their finances</b>, as they will be able to set their own prices and <b>manage their own inventory</b>. Finally, it can help them <b>build up their skills and knowledge</b> that can be used in other areas of business.</p>

45.	What help do people usually need when creating an e-commerce business?	<ol style="list-style-type: none"> <li><b>Business planning:</b> Developing a comprehensive business plan that outlines the vision, goals and strategies for the business is essential for success.</li> <li><b>Financing:</b> Securing startup capital to cover the costs of launching an e-commerce business can be difficult.</li> <li><b>Marketing:</b> A successful e-commerce business requires strong marketing efforts to reach customers and generate sales.</li> <li><b>Technology:</b> Building an online store requires a range of technology tools, including web hosting, shopping cart software and payment processing systems.</li> <li><b>Product sourcing:</b> Finding reliable suppliers for products is essential for any e-commerce business.</li> <li><b>Shipping &amp; logistics:</b> Developing an efficient shipping and logistics system is necessary to ensure orders are fulfilled quickly and accurately.</li> <li><b>Customer service:</b> Providing excellent customer service is a crucial part of any successful e-commerce business.</li> </ol>
46.	What are technical skills necessary for B40s to improve their lives?	<ol style="list-style-type: none"> <li>Digital literacy: Basic computer and internet skills, such as using a keyboard and mouse, navigating the web, and understanding different computer programs.</li> <li>Money management: Understanding how to budget, save and invest.</li> <li>Communication skills: Being able to effectively communicate both orally and in writing.</li> <li>Entrepreneurial skills: The ability to recognize and create opportunities, come up with creative solutions, and manage resources.</li> <li>Problem-solving skills: Being able to identify problems, develop strategies for solving them, and apply solutions to different situations.</li> <li>Leadership skills: Developing the ability to inspire and motivate others, build consensus among different stakeholders, and delegate tasks appropriately.</li> <li>Analytical thinking: Applying logical reasoning and critical thinking to understand complex situations and make decisions based on evidence.</li> <li>Networking skills: Developing relationships with people in your community who can provide support or resources that can help you improve your life.</li> </ol>
47.	What are products and services that can be offered with low skilled labor and low startup cost?	<ol style="list-style-type: none"> <li>Websites and Graphic Design: Offering basic website design and graphic design services requires minimal technical expertise and can be done with minimal startup costs.</li> <li>Virtual Assistant Services: Virtual assistants can provide a variety of services such as data entry, customer service, and bookkeeping, which require no specialized skills or technical expertise.</li> <li>Cleaning Services: Starting a residential or commercial cleaning business requires little capital investment and can be done with low-skilled labor.</li> <li>Pet Care Services: Offering pet sitting, pet walking, and pet grooming services requires no specialized skills or technical expertise, making it an ideal business to start with low-skilled labor and low startup costs.</li> <li>Home Repairs &amp; Maintenance: Home repair and maintenance services require minimal technical expertise and can be done with minimal startup costs.</li> <li>Food Delivery Services: Offering food delivery services requires no specialized skills or technical expertise, making it a great business to start with low-skilled labor and low startup costs.</li> </ol>
48.	What are the skills that need to be learnt by B40s to become self-reliant through e-commerce?	<ol style="list-style-type: none"> <li>Basic computer skills: B40s must learn how to use a computer and navigate the internet, including how to find and open webpages, type text, save files, and browse the web safely.</li> <li>E-commerce skills: B40s need to learn basic e-commerce skills, such as setting up an online store, creating product listings, managing customer accounts, and processing payments.</li> <li>Digital marketing skills: B40s need to learn how to promote their products or services online, including search engine optimization (SEO), pay-per-click (PPC) advertising, social media marketing, email marketing, and content marketing.</li> <li>Financial management skills: B40s should learn how to manage their money effectively, such as budgeting and keeping track of income and expenses.</li> <li>Customer service skills: B40s need to be able to provide excellent customer service in order to build a successful online business. This includes responding to customer inquiries in a timely manner and resolving customer complaints quickly.</li> </ol>
	Write an academic	<p><b>Introduction</b></p> <p>The B40s, or the bottom 40 percent of Malaysian households, are an important target group for economic development and poverty alleviation. Despite their low income, this group of people is increasingly turning to e-commerce as an alternative economic activity. This thesis investigates the opportunities and challenges that the B40s in Malaysia face in succeeding in e-commerce, and explores potential solutions to help them succeed.</p> <p><b>Background</b></p> <p>E-commerce has become a popular platform for economic activity in Malaysia, with the number of online shoppers increasing steadily in recent years. The B40s have been identified as an important target group for e-commerce, as they can benefit from improved access to goods and services at lower prices. However, there are a number of challenges that the B40s face when it comes to engaging in e-</p>

<p>thesis on the B40s in 49. Malaysia and how they can be assisted to succeed in e-commerce.</p>	<p>commerce. These include limited access to technology and digital literacy, financial constraints, lack of knowledge about e-commerce platforms, and limited access to customer service.</p> <p><b>Discussion</b></p> <p>In order to address the challenges faced by the B40s in engaging in e-commerce, a number of solutions have been proposed. These include government initiatives such as providing access to subsidized broadband services and digital literacy training, as well as private sector initiatives such as providing access to financing options and customer service support. Furthermore, there is potential for the use of technology such as artificial intelligence and machine learning to improve the user experience on e-commerce platforms, making them more accessible and easier to use.</p> <p><b>Conclusion</b></p> <p>The B40s in Malaysia have great potential to benefit from engaging in e-commerce. However, there are a number of challenges that they face which need to be addressed in order for them to succeed. Solutions such as government initiatives and private sector initiatives can help address these challenges, while advances in technology can improve the user experience on e-commerce platforms. With the right support, the B40s can gain greater access to economic opportunities through e-commerce.</p>
--	--

## 5. Conclusions

It is important for academic researchers to understand that GPT-3 is a “human-like text generator” that can generate plausible but possibly non-factual content. GPT-3 is also never designed to be factual, as it was trained on some sources that have questionable factuality (such as forums, etc.). Therefore, there is a **distressing need for fact-checking in GPT-3**. However, to GPT-3’s credit, it does seem to have some rudimentary knowledge base as it is trained on an extremely large dataset of text. Our experiments have polarly different questions indicated that **GPT-3 responses can be used to support whichever point of view that a researcher wants like to highlight with plausible explanations to support it**.

Consequently, the **risk for academic fraud** is extremely high with GPT-3. Although it can be a beneficial tool for learning if used properly, care must be taken so that GPT-3 responses are fact checked so that is not abused by researchers and students to write entire assignments or research papers. Another angle to consider is GPT-3 could be used to generate **dubious academic papers** that the public trusts. Our experiments indicate that GPT-3 already has some **rudimentary understanding of the structure of an academic paper** and **can be misused by individuals that lack the ethical and cognitive ability to use this tool judiciously**.

Finally, based on our experiments, we postulate that **GPT-3 can be used to help academic writers**, but the tool **must be used with care**. It can be used to form beginning ideas, or to create introductory parts of the paper due to its ability to describe current knowledge consensus. This is because plausible information is relatively neutral and can ease the reader into more technical aspects of the paper. However, care must be taken to **fact-check its statements carefully against research in the area**. Another way is to train GPT-3 to **cite the sources for its material as well**, perhaps in combination with existing websites such as Google Scholar, Scopus and IEEExplore.

## 6. What should be GPT-3’s future direction?

GPT-3 is truly a remarkable invention with limitless potential for humankind. As such, we propose several future directions for OpenAI engineers to consider as GPT-3 has the potential to replace search engines like Google because of its NLP capabilities:

1. The authors are advocating that the next generations of the GPT-3 be **trained with data from academic research papers**. GPT-3 was **not trained to source information from academic research papers in PDF format** the format commonly used to disseminate research papers. This strategy could be an invaluable addition to GPT-3's knowledge base as academic research is factual, evidence-based and peer-reviewed. This could be used to **counterbalance GPT-3's possible information bias**, as it admits in Question 11.
2. The authors believe that the inclusion of academic papers can help **defeat many false news and claims** made by unscrupulous individuals that have significant social media presence and influence on the internet, and thus urge OpenAI to consider these modifications to future versions of GPT-3. The implementation can perhaps in the form of a slider to adjust between "Casual Mode" to "Factual Mode".
3. Consequently, we can also consider two versions of GPT-3, **one trained on factual data, and the other trained on other types of information**. Some researchers have already worked on compiling a high-quality text corpus targeted at training large-scale language models derived from academic and professional sources [26], which could be integrated as part of GPT-3's knowledge base.
4. In this multimedia age, some useful information can also be gained from online video platforms. GPT-3 could be combined with the **OpenAI Whisper** project to **transcribe videos** and **extract information** from them.
5. **GPT-3 was not trained on the Dark Web**. The authors tend to agree on OpenAI engineers' decision of this. However, it would be interesting to examine the outcome of this as although the Dark Web contains questionable content, it is still a rich source of information, even larger than the information stored on the surface web. The authors invite further discussion on this matter to further enhance and complete GPT-3's knowledge based on its pursuit for the perfect NLP model.

## Acknowledgements

The authors would like to thank the Ministry of Education, Malaysia for funding this research through the Long-Term Research Grant (LRGS) grant (Grant number 600-RMC/LRGS 5/3 (001/2020), Project Code: 251905210001).

## References

1. <sup>a, b, c, d</sup>C. Si et al., "Prompting GPT-3 To Be Reliable," Oct. 2022.
2. <sup>^</sup>S. Saravanan and K. Sudha, "GPT-3 Powered System for Content Generation and Transformation," in *2022 Fifth International Conference on Computational Intelligence and Communication Technologies (CCICT)*, Jul. 2022, pp. 514–519. doi: 10.1109/CCICT56684.2022.00096.
3. <sup>a, b, c, d, e</sup>S. Nath, A. Marie, S. Ellershaw, E. Korot, and P. A. Keane, "New meaning for NLP: the trials and tribulations



- of natural language processing with GPT-3 in ophthalmology,” *British Journal of Ophthalmology*, vol. 106, no. 7, pp. 889–892, Jul. 2022, doi: 10.1136/bjophthalmol-2022-321141.
4. <sup>a</sup>H.-Y. Lin, “Large-Scale Artificial Intelligence Models,” *Computer (Long Beach Calif)*, vol. 55, no. 5, pp. 76–80, May 2022, doi: 10.1109/MC.2022.3151419.
  5. <sup>a</sup>C. Edwards, “The best of NLP,” *Commun ACM*, vol. 64, no. 4, pp. 9–11, Apr. 2021, doi: 10.1145/3449049.
  6. <sup>a</sup>B. Kim et al., “What Changes Can Large-scale Language Models Bring? Intensive Study on HyperCLOVA: Billions-scale Korean Generative Pretrained Transformers,” in *Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing, 2021*, pp. 3405–3424. doi: 10.18653/v1/2021.emnlp-main.274.
  7. <sup>a</sup>R. Lowe, N. Pow, I. Serban, and J. Pineau, “The Ubuntu Dialogue Corpus: A Large Dataset for Research in Unstructured Multi-Turn Dialogue Systems,” in *Proceedings of the 16th Annual Meeting of the Special Interest Group on Discourse and Dialogue, 2015*, pp. 285–294. doi: 10.18653/v1/W15-4640.
  8. <sup>a</sup>R. E. Wray, J. R. Kirk, and J. E. Laird, “Language Models as a Knowledge Source for Cognitive Agents.”
  9. <sup>a, b</sup>R. Dale, “GPT-3: What’s it good for?,” *Nat Lang Eng*, vol. 27, no. 1, pp. 113–118, Jan. 2021, doi: 10.1017/S1351324920000601.
  10. <sup>a, b</sup>K. McGuffie and A. Newhouse, “The Radicalization Risks of GPT-3 and Advanced Neural Language Models,” *Sep. 2020*.
  11. <sup>a, b, c, d</sup>A. Sobieszek and T. Price, “Playing Games with Ais: The Limits of GPT-3 and Similar Large Language Models,” *Minds Mach (Dordr)*, vol. 32, no. 2, pp. 341–364, Jun. 2022, doi: 10.1007/s11023-022-09602-0.
  12. <sup>a, b</sup>D. Su et al., “Context Generation Improves Open Domain Question Answering,” *Oct. 2022*.
  13. <sup>a, b, c</sup>Chenglei Si, Naman Molri, Gurmehar Cheema, Elliot Huang, and Arjun Akkiraju, “Benchmarking GPT-3 For Closed-Book QA: Strengths and Weaknesses,” *Maryland, 2022*.
  14. <sup>a, b, c</sup>T. Goyal, J. J. Li, and G. Durrett, “News Summarization and Evaluation in the Era of GPT-3,” *Sep. 2022*.
  15. <sup>a, b</sup>A. Bhaskar, A. R. Fabbri, and G. Durrett, “Zero-Shot Opinion Summarization with GPT-3,” *Nov. 2022*.
  16. <sup>a</sup>G. Spitale, N. Biller-Andorno, and F. Germani, “AI model GPT-3 (dis) informs us better than humans,” *Jan. 2023*.
  17. <sup>a, b</sup>M. Ravallion, “Ethnic inequality and poverty in Malaysia since May 1969. Part 2: Poverty,” *World Dev*, vol. 134, p. 105039, Oct. 2020, doi: 10.1016/j.worlddev.2020.105039.
  18. <sup>a</sup>H.-A. Lee and M. A. Khalid, “Is Inequality Really Declining in Malaysia?,” *J Contemp Asia*, vol. 50, no. 1, pp. 14–35, Jan. 2020, doi: 10.1080/00472336.2018.1551560.
  19. <sup>a</sup>Goh Lim Thye, Selvarajan Sonia Kumari, Yong Sook Lu, and Shahabudin Sharifah Muhairah, “The Impacts of External and Internal Uncertainties on Income Inequality in The ASEAN 5 Countries,” *Int. Journal of Economics and Management*, vol. 15, no. 3, pp. 351–375, 2021.
  20. <sup>a</sup>P. V. Kebangsaan Ekonomi Malaysia ke, E. Berpendapatan Tinggi, T. ke Arah Peningkatan Inovasi, P. dan Kualiti Hidup, and M. Abdul Khalid, “NEP to NEM: Who Cares? Wealth Distribution in Malaysia,” 2011.
  21. <sup>a</sup>Z. Shafi, “ETHNIC HETEROGENITY IN THE MALAYSIAN ECONOMY,” 2009.
  22. <sup>a</sup>S. Anand, “ASPECTS OF POVERTY IN MALAYSIA.”
  23. <sup>a, b, c, d, e, f</sup>A. H. Roslan, “Income Inequality, Poverty and Development Policy in Malaysia.”
  24. <sup>a</sup>N. J. White, *British Business in Post-Colonial Malaysia, 1957-70*. Routledge, 2004. doi: 10.4324/9780203333020.

25. <sup>^</sup>M. Ali Embi dan Norapiah Abd Rahman, “The Impact of Equal Employment Opportunity-Affirmative Action... The Impact of Equal Employment Opportunity-Affirmative Action in U.S. and New Economic Policy in Malaysia on Employment and Reverse Discrimination,” *Jurnal Demokrasi*, vol. 8, no. 2, 2009.
26. <sup>^</sup>L. Gao et al., “The Pile: An 800GB Dataset of Diverse Text for Language Modeling,” Dec. 2020.