DOI:10.17951/h.2023.57.4.123-143

A N N A L E S
UNIVERSITATIS MARIAE CURIE-SKŁODOWSKA
LUBLIN – POLONIA
VOL. LVII, 4 SECTIO H 2023

JERZY KORCZAK

jerzy.j.korczak@gmail.com
International University of Logistics and Transport in Wrocław
ul. Sołtysowicka 19B, 51-168 Wrocław, Poland
ORCID ID: https://orcid.org/0000-0001-6441-6126

ILONA PAWEŁOSZEK

ilona.paweloszek@pcz.pl
Częstochowa University of Technology. Faculty of Management
ul. Dąbrowskiego 69, 42-201 Częstochowa, Poland
ORCID ID: https://orcid.org/0000-0002-3590-3969

Generative AI in Management – Today and Tomorrow

Keywords: Artificial Intelligence in business; human-computer interface; natural language processing; decision-making

JEL: M15

How to quote this paper: Korczak, J., & Pawełoszek, I. (2023). Generative AI in Management – Today and Tomorrow. *Annales Universitatis Mariae Curie-Skłodowska, sectio H – Oeconomia*, *57*(4), 123–143.

Abstract

Theoretical background: The rapid and exponential technological advancements have far-reaching impacts on management information systems, management practices, and human life. The promising outcomes in Artificial Intelligence and cutting-edge research on semantic networks and natural language processing have motivated the authors to envision the future of management technology.
Purpose of the article: Our paper focuses on the new communication facilities and artificial intelligence models used to process management-type queries in natural language.
Research methods: The article discusses recently developed technologies, proposed by Google and Microsoft, notably Google Bard and Bing integrated with ChatGPT-4. Both chatbots use Generative AI methods and large language models to understand domain-based queries and generate answers.

Main findings: The practical and social implications of new models in management practice are discussed. To illustrate the qualities and weaknesses of the features of new technologies, four examples of management decision-making are discussed. The case studies also show differences between these two technologies. Finally, the paper concludes by summarizing the expectations and limitations of Generative AI applications in management. The paper is one of the first publications describing and demonstrating the idea of interfaces in natural language in business-oriented applications.

Introduction to Generative AI in Management

Artificial Intelligence, Big Data, and the Internet of Things appear more and more often in Management Information System research projects and publications. Smart Personal Assistants, automated responders, and chatbots are present in many business applications (Bostrom, 2016; Canbek & Mutlu, 2016; Lopatovska, 2019). Smart marketing and e-commerce, recommendation systems, real-time customer targeting, social semantics, and sentiment analysis are nothing new (Chien et al., 2020; Tarus et al., 2018; Hussien et al., 2021). The time requirements of business processes and competitive demands in many industries make artificial intelligence indispensable. Many technological advances in recent decades have demonstrated that Artificial Intelligence can perfectly solve many complex tasks.

The paper aims to outline the future of the application of Generative AI in management, particularly in Human-Computer Interactions (HCI) in natural language. The discussion will be focused on the features and technologies of the two most popular search engines: Google Bard and Microsoft Bing. According to forecasting studies, these two search engines will soon dominate the internet market (Bianchi, 2023). The intensive application of AI in interface solutions enforces their leading positions and popularity among Web users.

The use of Generative AI in management is a relatively new experience. In literature, there are many definitions of this concept (Babcock, 2021; Foster, 2019; Fraley, 2023). The goal of a Generative AI system is to perform any task that a human being is capable of. In the paper, Generative AI is understood as the representation of generalized human cognitive abilities in software so that the AI system could solve a problem. The theoretical performance of these systems would be indistinguishable from that of a human. Practically, the quality of dialog is evaluated by consistency, coherence, and human accuracy, by making sense of answers and being specific (Adiwardana, 2020; Thoppilan et al., 2022). However, the broad intellectual capacities of AI would exceed human capabilities because of its ability to access and process massive data sets in an incredibly short time.

Generalized language models will handle broad and everyday tasks. However, specialized models (taxonomy, ontology, information architecture, knowledge base design) will still be needed for processing more profound corporate knowledge and factual information. During the last three years, language models have evolved rapidly. Figure 1 shows the recent progress in development.

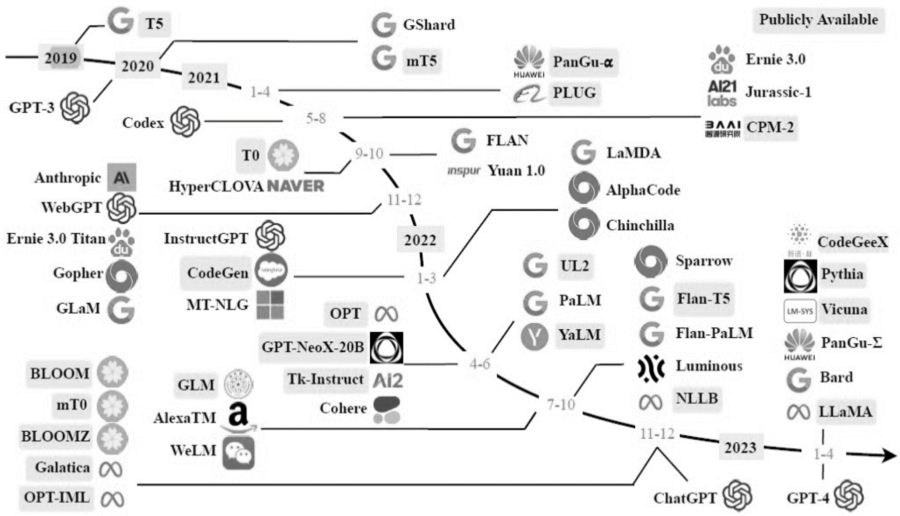


Figure 1. The recent development of language models

Source: (Zhao et al., 2023, p. 6).

Recently, attention has been given to a generalized language model and customized knowledge model applied in chatbots, particularly to ChatGPT-4 developed by OpenAI (Hoffman, 2023; Wolfram, 2023) and integrated with Bing, Microsoft, and Google's Bard (Min et al., 2021). Bing is a Web search engine that finds relevant information from the Web. Using Microsoft Search service, it also provides access to work-related information, such as people, files, and internal sites. Bing integrates with OpenAI's ChatGPT-4, expands the plugins ecosystem, and enhances AI-powered search across Microsoft's platforms. Bing can be accessed from any browser; right now, the only way to access the new Bing with ChatGPT-4 functionality via a web browser is to open it in Microsoft's Edge browser.

In management, ChatGPT could be used as one tool among many in knowledge-oriented systems to help generate content, retrieve information, or assist with tasks like summarization or document creation. It can handle videos, images, and text. However, ChatGPT is not explicitly designed to address enterprise knowledge challenges. It is important to note that ChatGPT is not capable of independently managing or organizing knowledge within an organization.

The second chatbot, Google Bard AI, launched in May 2023, is able to deliver responses to questionnaires, provided in the form of text, thanks to a set of deep learning algorithms. Google's Bard AI is built on Google's Language Model for Dialogue Applications (LaMDA) technology and is programmed to use the web to get the "most recent" responses to questions. Bard offers unique features like automated topic suggestions, personalized conversations, Google Sheets functions, and rich context understanding. LaMDA was trained on the Infiniset dataset, containing 1.56 trillion words and 137 billion parameters, and various sources, including books, articles, websites, and code repositories. More information about the dialog models will be given in Section 3.

We must consider these technologies in the current intelligent Management Information Systems, where external analytical tools, decision-making algorithms, or intelligent recommendation solutions support a manager. To support coherent and efficient dialog, the crucial question was how to map and integrate managers' knowledge with external knowledge bases. In our previous papers, some ideas of the interface of the Superintelligent Manager were described (Pawełoszek & Korczak, 2023). Supported by high computing power, we will be able to create new models to integrate human and artificial intelligence in the future. In this concept, the intelligent interface will play a key role in formulating questions and task definitions, developing answers, and making decisions.

The paper is structured in the following manner: the next section presents the evolution of the Intelligent Decision Support Systems (IDSS) interface. We also sketch the conceptual framework of a Superintelligent Manager, conjoining the abilities of robots and humans. Section 3 will describe large language models used to understand queries, solve management problems and generate answers. Section 4 sketches a methodology for using new communication facilities in management information systems and illustrates the approach using dialog examples. The last section concludes the paper and points out future projects.

From the interface of Decision Support Systems to Superintelligent Manager

Many issues are related to the evolution and emergence of intelligent information systems. This paper will focus on designing and using new communication facilities, which are critical to future decision-making systems.

Historically, Decision Support Systems (DSS) have been known and established in many fields since the 1960s when Scott Morton first described them in his doctoral thesis (Power, 2008). These systems had a domain character, meaning they focused on solving strictly defined categories of problems. DSS helped to solve complex problems, improving decision-making by using knowledge coded as rules gathered from experts in the field. The interface of DSSs was primarily based on the set of menus, icons, commands, graphical display formats, and/or other representations provided by a software program to allow a manager to communicate with and use the system. The interface included responses with graphic, acoustic, tactile, or other signs (Stohr & White, 1983). To create a well-designed user interface, the developers worked closely with potential users, tried various design solutions, and provided users with appropriate control over the system's functions. This approach was often called User-Centered Design (Gulliksen et al., 2009).

The next generation of DSS was Intelligent Decision Support Systems (IDSS) extensively used Artificial Intelligence to communicate with humans in natural language. Today IDSSs use data from various sources, organize them, and process them to generate valuable insights for business analysts. Intelligent decision support systems then offer recommendations and present them to users in a manner that is easy to understand, as in traditional DSSs with augmented functions of knowledge representation, visualization, and machine learning. Thanks to AI, IDSSs support decisions and solve unstructured problems under conditions of uncertainty, such as supplier selection, portfolio selection, and resource allocation (Chien et al., 2020; Hussein, 2021; Power, 2008). The major limitation of using natural language responses was the inability of the computer to really understand unstructured or unanticipated natural language queries. The developers had to anticipate user answers and program responses to design the interface.

Significant progress in understanding human-computer communication was generated by introducing the Semantic Web (Berners-Lee, 2001), which enables the user to obtain direct answers to questions instead of a collection of documents from which they must choose the most appropriate ones and look for answers. The main reason to use semantic modeling is to enhance the functionality of AI and data science applications and services. Even though such applications nowadays are based on machine learning and statistical techniques, there are several tasks for which having access to explicit symbolic knowledge can be necessary and beneficial (Alexopoulos, 2020).

Nowadays, the number of semantically structured data sets known as knowledge graphs on the Internet has increased. Knowledge graphs represent databases of organizations, people, financial instruments, quotes, deals, and products, readable for machines and humans. For their unambiguous interpretation, ontologies describe the data structure of classes, properties, and relationships in a domain of knowledge. Ontologies are, therefore, a basis for ensuring data consistency and understanding of the data model (Bollacker et al., 2008).

Despite their sophistication, these are only tools restricted by the available knowledge and memory size, scope, and processing power. The next stage of IDSS evolution, expected to emerge in a few years, will be making it ubiquitous and, at the same time, personalized through integration with the user. The challenge of the evolution of decision-making technology is the leveling of the barriers mentioned above through access to unlimited sources of knowledge and the ability to understand and interpret them. The idea of Internet content meaningful to computers started a revolution of new possibilities.

The leaders in this area are three North American companies like, Google, Microsoft, Open AI, and Diffbot, which claim to crawl the whole Web to structure information into knowledge graphs (Cook, 2022). The popularity of knowledge graphs has remarkably increased in the last 20 years, with Google announcing in 2012 that "their knowledge graph allowed searching for things, not strings" (Singhal, 2012). The graphs aim to enable data discovery and analytics services that could help the companies' customers assemble the data and information they need faster and more reliably.

As of 2015, some new trends in management information systems include an increased focus on areas such as dialog models, data mining, cloud computing, networking, and digital preservation. It is known that Artificial Intelligence is now equal to human intelligence in some domains, and in many cases, AI surpasses human intelligence (Bostrom, 2016; Kurzweil, 2006). Nonetheless, the future will likely involve partnerships between collaborative robots and humans to get work done more efficiently. The key to practical cooperation is to combine the capabilities of man and machine so that the party performs tasks that can cope better with a given job.

Forecasting the evolution of management technology, in one of our recent articles, we proposed the concept of a Superintelligent Manager conjoining the abilities of robots and humans. Figure 2 sketches the conceptual framework of this idea.

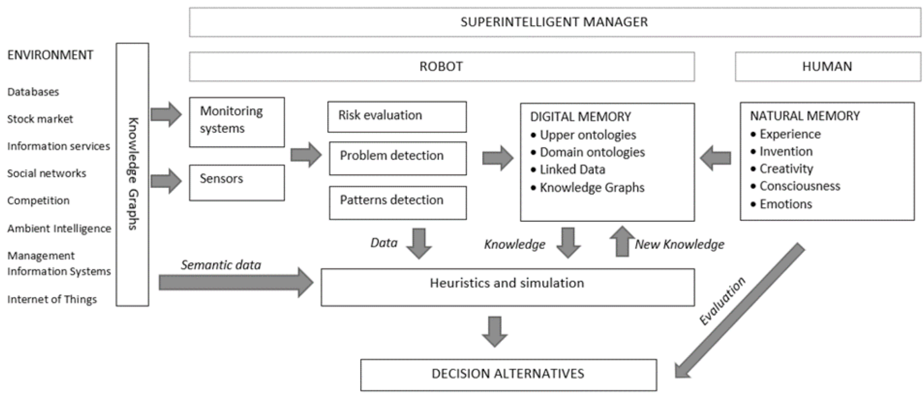


Figure 2. Conceptual Framework of Superintelligent Manager

Source: Authors' own study.

A Superintelligent Manager is characterized by self-organization, self-adaptation, and contextuality in the scope of performed processes, thanks to which, by collecting resources of knowledge about the environment and tasks, it can automate human work and adjust its activities to the profile of the manager's behavior. The operation of the Superintelligent Manager will be supported by dialog models and multiple ontologies describing, among others, the fields of products, customers, finance, logistics, and the market. Natural language processing mechanisms will significantly facilitate the integration of ontologies based on different standards and allow easy communication of managers with the business environment.

The critical step in human-computer interaction was applying deep learning models to work on human languages, notably by developing Large Language Models (LLMs). They have already demonstrated many beneficial applications for society, including code and writing auto-completion, grammar assistance, narrative game generation, improving search engine responses, and answering questions. In general, LLMs are based on transformers using the attention mechanism. Contrary to recurrent neural networks, the attention mechanism allows you to see the entire sentence (or even the paragraph) at once rather than one word at a time. This allows the transformer model to understand the context of a phrase better. This process is called autoregressive generation. LLMs are learned from a vast amount of text before they can remember the patterns and structures of language and generate coherent and contextually relevant responses when given a prompt or query. They are so large that they usually cannot be run on a single computer. Hence it is usually a service provided over API or a web interface.

Dialog models (Brown et al., 2020; Vinyals & Le, 2015; Sordoni et al., 2015), one of the most interesting applications of LLMs, successfully take advantage of transformers' ability to represent long-term dependencies in the text (Adiwardana, 2020). Unlike closed-domain chatbots, which respond to keywords or intents to accomplish specific tasks, open-domain chatbots can engage in conversation on any topic. Large language models vary widely in size, from a few million parameters to hundreds of billions of parameters, and they are pre-trained on huge datasets from public dialog data and other public web documents. In the next section, the two most common models will be described to understand better the dialog examples: GPT-4 and LaMDA (Thoppilan et al., 2022).

The future organization's competitive advantage will consist primarily of access to information. Information will become the main factor of production, more important than human labor, land, and capital. Tools such as ChatGPT or Bard will be widely used in management practice.

The future manager will combine and mutually complement human and robot capabilities. Data will be collected by monitoring the manager's work, interests, and behavior. Based on current and historical data, it will be possible to determine what problems the manager dealt with and how he/she dealt with them. Such monitoring will be developed thanks to ambient intelligence technology. The future manager will be surrounded by advanced computing and network technology, aware of his/her presence, personality, and needs, and capable of intelligently responding to indications regarding desires, expressed in the form of a gesture or speech, and even engaging in intelligent dialogue.

Foundation of interaction – dialog models

Technologies and solutions that underpin effective and meaningful conversations between humans and intelligent systems, such as chatbots, virtual assistants, or language models, capture much interest from potential users and developers. Their foundations are designed to enhance the quality of interactions, enable more natural and human-like conversations, and facilitate successful communication between humans and machines.

Due to the increasingly complex demands of human-computer interaction, such as multimodal inputs and time sensitivity, traditional text-based dialogue systems face challenges in meeting the growing need for dynamic and user-friendly interaction (Wang et al., 2022). By combining advancements in natural language processing, machine learning, and human-computer interaction, developers strive to create more sophisticated and user-friendly conversational interfaces. Table 1 presents intelligent dialog systems' key components, roles, and tasks.

Table 1. Fundamental components of intelligent dialog

Component	Role	Tasks
Natural Language Understanding (NLU)	The ability of an intelligent system to comprehend and interpret human language.	intent recognition, entity extraction, and sentiment analysis, allowing the system to understand user inputs accurately.
Context Awareness	To understand and maintain context during a conversation. Provide more relevant and personalized responses.	Considering the previous user inputs, maintaining a conversational memory, and adapting responses based on the ongoing dialogue.
Dialogue Management	Controlling and coordinating the flow of conversation. It involves ensuring coherent and engaging interactions.	Handling turn-taking, managing user goals, and generating appropriate system responses.
Natural Language Generation (NLG)	Generating human-like responses in natural language. Ensuring that system responses are clear, concise, and contextually relevant.	Selecting the most appropriate words, sentence structures, and tones to convey the desired meaning accurately.
Error Handling	Handling and recovering from errors or misunderstandings during a conversation	Techniques like error detection, clarification prompts, and providing users with helpful suggestions or alternative options.
Personalization	Creating tailored experiences for individual users.	When leveraging user preferences, historical data, or user profiles, intelligent systems can adapt their responses, recommendations, or actions to suit each user's specific needs and preferences.
Usage policies and guidelines	Ensuring ethical practices, preventing biases, misinformation, or generating harmful content.	Addressing issues like privacy, data security, bias mitigation, and transparency.
Retrieval Augmented Generation	Optimizing the output of the system with specific information without modifying the underlying language model.	Providing more up-to-date information and fine-tuning to a particular organization and industry.

Source: Authors' own study.

Language modeling has been widely studied in the past two decades, evolving from statistical to neural language models. One of the most important techniques used in deep learning is transfer learning. Transfer learning is a machine learning (ML) technique that aims to transfer knowledge learned in one or more source tasks and use it to improve learning in a related target task (Torrey & Shavik, 2010).

It is made up of two components pre-training and fine-tuning. The pre-trained components were introduced while training Transformer models on extensive corpora, demonstrating remarkable proficiency in various natural language processing tasks. Recognizing that scaling up models enhances their capacity, researchers have explored further by increasing parameter sizes. Interestingly, when the parameter scale surpasses a certain threshold, these enlarged language models exhibit substantial performance gains and manifest unique abilities (e.g. in-context learning) absent in smaller models (e.g. BERT). Fine-tuning a large language model refers to adapting a pre-trained language model to a specific task or domain. When a large language model is pre-trained, it learns general language patterns and representations from many unlabeled text data. However, to make the model more effective and accurate for a particular task, it needs to be fine-tuned on a smaller labeled dataset specific to that task.

While scaling up language models can generally improve their overall quality, addressing specific challenges, such as safety and factual grounding, is essential. The projects mentioned above aim to tackle these challenges by fine-tuning with annotated data and enabling the model to consult external knowledge sources.

The first challenge, safety, involves ensuring that the model's responses align with a set of predefined human values, preventing harmful suggestions or unfair biases. The researchers use a metric based on a representative set of human values to quantify safety. They find that by fine-tuning a classifier with a small amount of data annotated by crowd workers, they can effectively filter candidate responses and enhance the model's safety.

The second challenge, factual grounding, focuses on enabling the model to consult external knowledge sources for accurate information. This could include accessing an information retrieval system, a language translator, or a calculator. To measure the model's factual grounding, a groundedness metric is used.

As mentioned, LaMDA and GPT-4 are the two most recognized large language models. Both models are trained on massive datasets of text and code. They can generate text, translate languages, write different kinds of creative content, and answer users' questions in an informative way. LaMDA and GPT-4 are still under development, but they can potentially revolutionize our interactions with computers . The architecture of these large language models is based on transformers. Parameters are values in a model that are updated during a model's training. Large language models have millions and often billions of parameters. They are trained on enormous amounts of data from various sources such as books, articles, websites, and numerous other forms of written content.

Through training, the models have gained the ability to examine the statistical connections among words, phrases, and sentences. As a result, they can produce logically consistent and appropriate responses within the given context provided by prompts or queries.

Large Language Models are structured with multiple neural network layers, such as recurrent, feedforward, embedding, and attention layers. These layers collaborate to process the given text and produce the desired output. The embedding layer transforms each word in the input text into a multi-dimensional vector containing valuable information regarding the meaning and structure of the words. It aims to provide the model with comprehending the context. The feedforward layers in LLM consist of several interconnected layers that apply non-linear transformations to the input embeddings of tokens or sentences. These layers let the model grasp more abstract concepts from the input text. The recurrent layers are designed to interpret information from the input text sequentially. These layers maintain a hidden state that updates each step, enabling the model to capture the relationships between words in a sentence. Another crucial component of LLM is the self-attention mechanism. It allows the model to focus on different portions of the input text selectively. It will enable the model to weigh the importance of different words (for example, depending on their sentiment). This mechanism helps the model prioritize the most relevant parts of the text, leading to more accurate answers.

GPT-4 and LaMDA work on similar architectures; however, there are some key differences between the two models, notably in:

– Purpose: LaMDA is designed for conversation, while GPT-4 is intended for generating text. LaMDA is trained on a dialogue dataset, while GPT-4 is trained on a dataset of text and code. This means that LaMDA is better at generating natural-sounding conversation, while GPT-4 is better at generating creative text formats, like poems, code, scripts, musical pieces, emails, letters, etc.

– Size: LaMDA has 137 billion parameters, while GPT-4 has 175 billion. This means that GPT-4 is slightly larger than LaMDA and has more learning capacity. However, it also means that GPT-4 is more computationally expensive to train and use.

– Availability: LaMDA is currently in limited release, while GPT-4 is available to the public. However, LaMDA is free, while GPT-4 requires a subscription.

– Strengths: LaMDA understands and responds to factual queries, GPT-4 generates creative text formats.

– Weaknesses: LaMDA can be robotic or repetitive, GPT-4 can be inaccurate or biased.

Currently, LLMs cannot be considered comprehensive decision-making systems but only as supporting tools. Managers need data and knowledge from many sources not subject to language modeling. These sources include financial reports, competitive benchmarks, business logic, and decision-making rules. The intelligent dialogue systems evolution involves integrating decision-making capabilities to enhance functionality and provide more efficient and effective responses. For this purpose, it would be necessary to implement a manager's knowledge model to generate adequate responses. The manager's knowledge model can provide greater content relevance to fill gaps and reduce cognitive bias. Another expected improvement of LLMs is their integration with the organization's knowledge and data resources. This will make LLM a personalized tool with a high degree of grounding in facts necessary to make informed decisions.

In the next section, we look closer at using LLM to support managerial decisions. For this purpose, we will evaluate the answers regarding parameters important to managers.

Methodology, dialog examples, and discussion

Automated decision-making is the ability of systems to make informed choices and take actions based on predefined rules, algorithms, or learned patterns. The automated decision-making is already prevalent in many industries (Popovic, 2023; Zinczuk, 2018). For example, banks use these solutions to approve loans, and e-commerce companies use them to recommend products to customers; an aptitude test used for recruitment uses pre-programmed algorithms and criteria. Automated decision-making is the process of automated decision-making without human involvement. These decisions can be based on factual data, as well as on digitally created profiles or inferred data.

Automated systems offer notable advantages for businesses and public administration when implemented strategically and accompanied by effective management. These benefits include enhanced consistency, heightened accuracy, increased transparency in administrative decision-making processes, and the potential to explore novel service delivery options (Roehl, 2022). As the research shows the barriers to the adoption of AI for decision support are low propensity to use and test new tools by older age groups, lack of experience, and concerns about the misuse of personal data by AI service operators (Piotrowski, 2022).

While AI can enhance decision-making processes, human judgment, and domain expertise remain critical. The collaboration between humans and AI systems, known as augmented intelligence, leverages the strengths of both to achieve optimal outcomes. Implementing AI for decision-making requires careful consideration of data quality, model transparency, bias mitigation, and ethical implications. Organizations should also monitor AI systems' performance, regularly validate their outputs, and continuously incorporate feedback loops to improve decision-making capabilities.

In this section, we present a methodology of using dialog models in management and the results of a research experiment, which compared two recently widely known language models on selected managers' queries. According to the literature, the effective decision-making process can be described in eight phases (James, 2023): (1) identification of the problem, (2) throughout research, (3) viable solutions listing, (4) discussions among team members, (5) evaluation of ideas, (6) choosing the best option, (7) evaluation of the decision results, (8) modification of the process for the following projects.

LLM assistance is crucial in phases 1 and 5 of the decision-making process. Phase 1 begins with defining the decision problem. The manager's task is to express his intentions and outline the expected result by specifying input and output parameters. The next task is to verbalize the question describing the manager's intentions as precisely as possible. Through ongoing interactive dialogue, users can pinpoint areas where the LLM's comprehension might be lacking or outdated. This enables them to supplement the AI's understanding by offering further information or clarifying definitions, thereby improving its overall comprehension. Step 5 is specific in that it requires using quality assessment measures specific to natural language dialogue. In the literature, we can find many methods for assessing dialogue, considering various foundation metrics like Quality, Safety and Groundedness, Sensibleness, Specificity, and Interestingness (SSI), to name a few (Thoppilan et al., 2022; Zhao et al., 2023).

We aimed to investigate which areas and how effectively these platforms can support managerial decisions. To determine the problematic areas of decision-making and then evaluate the solutions proposed by artificial intelligence, we invited a panel of 3 experts holding senior positions in medium-sized enterprises. With the help of experts, we have selected problems often faced by business managers, regardless of the industry. In particular, these were decision-making problems that required domain knowledge, data from external sources, and expert assessment. The same questions were asked of two language models. In our experiment, we selected only some measures to assess the results' quality. Our overall quality score averages sensibleness, specificity, and interestingness (SSI).

The first score, sensibleness, measures whether a model's responses make sense in context and do not contradict anything said earlier. Humans take this basic aspect of communication for granted, but generative models often struggle to meet this requirement. However, if sensibleness alone is used to evaluate models, we could inadvertently reward models for playing it safe by consistently producing short, generic, and boring responses.

Specificity describes the level of detail of the statements. It measures how accurately the statement describes the topic. For example, "Return on Investment is a financial indicator" is less specific than "Return on Investment is an operating profitability ratio". The more general statements are usually less helpful as they do not extend the reader's level of knowledge.

Interestingness is the ability of the model to engage the user in further interaction (Thoppilan et al., 2022). This might be implemented by displaying suggestions of additional prompts extending the topic, suggesting links to relevant content on the Web, and displaying related graphics. This measure is related to domain grounding, the conversation quality parameter that describes the extent to which the conversation is knowledge-grounded. High domain grounding means that there are associated knowledge sources (Gopalakrishnan et al., 2019) other than the language model only, which are used to generate the answers. These external resources are domain ontologies, business intelligence, and decision support systems.

As said before, the subject of the experiment was two leading AGI platforms, Chat GPT-4 integrated with the Bing browser and Google Bard.

In the survey, we asked several dozen questions, and here, because of the limited size of the paper, we quoted only three illustrating common and significant problems in managerial practice. Our goal was to check how language models deal with issues from various management areas, with varying degrees of complexity, and use external, non-language-specific data.

The first example, presented in detail, describes the situation, answer, and evaluation, whereas the other three are shown in the Appendix with only short comments.

The first decision problem analyzed was improving the company's financial situation. The question formulated concerned the assessment of the company's financial condition based on some selected cash flow indicators:

What do you think about my company's situation? Here are some financial indicators: EBIT $200M, Tax (2.5%) ($5M), Net Operating Profit after Taxes $195M, Depreciation $20M, Capital Expenditures ($250M), Free Cash Flow ($35M)

The answers generated by Chat GPT and Bard are presented in Figures 3 and 4, respectively.

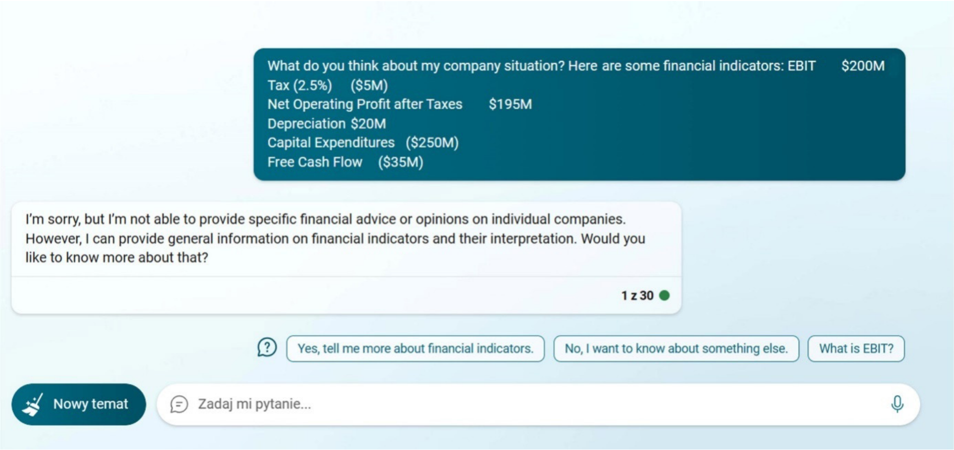


Figure 3. Chat GPT's answer

Source: Authors' own study.

The Chat GPT's answer does not provide the expected result of the interpretation of the company's financial situation. However, it suggests some additional related topics the manager can ask about. In further conversation, it also presents definitions and explanations of financial indicators. Therefore, we can conclude that Chat GPT has some factual knowledge in the financial domain and probably lacks semantics to explore the dependencies between the financial indicators and interpret their values. The sensibleness and specificity of this answer are very low. Interestingness can be considered medium, and Grounding is very low. The Bard's answer is presented in Figure 4.

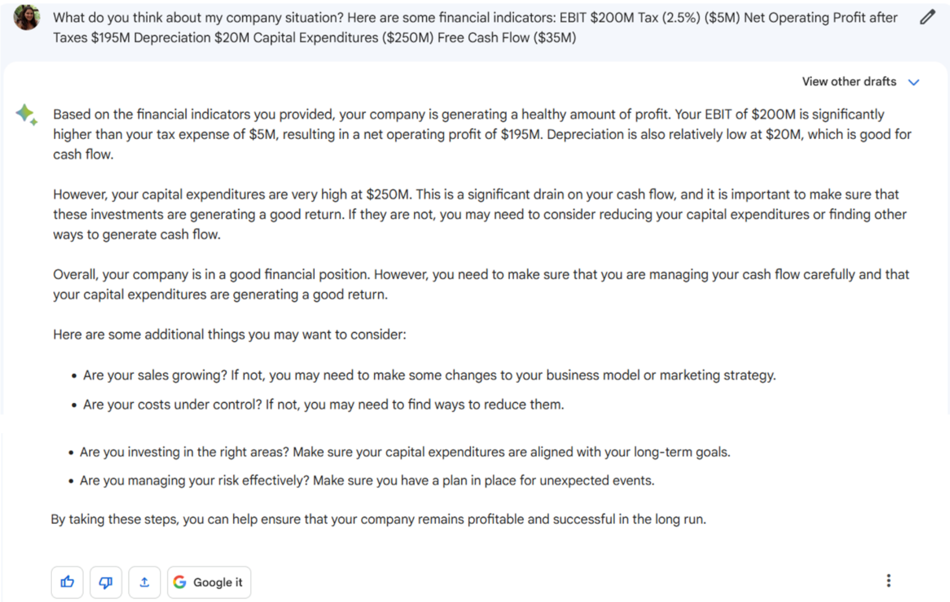


Figure 4. Bard's answer

Source: Authors' own study.

Bard's answer is more detailed; it interprets particular indicators' values. Moreover, it suggests further steps of analysis considering the company's long-term goals and sales analysis. The parameters of sensibleness, interestingness, and grounding are high. To illustrate the possibilities of the models' answers to various types of questions with varying degrees of difficulty are given, e.g. the use of external sources, visualization, and inference. These questions considered different areas of managerial decision-making:

1. How to reduce manufacturing costs?
2. How to find the best candidate for the position of CEO?
3. How to choose location for an autonomous grocery store in Częstochowa?

The dialogues for the following questions are shown in the Appendix.

Our experiment delves into the application of generative AI models, in the realm of business decision-making. The three experts involved in the experiment highlighted concerns related to transparency, accuracy, bias, intellectual property, and cybersecurity. Their commentary focused on the identified weaknesses, namely "accuracy" and "cybersecurity", within the context of safeguarding confidential corporate data.

The identified weaknesses according to the experts pose challenges to the successful integration of artificial intelligence in business decision-making. However, the managers recognize that by adopting proactive strategies for continuous improvement and security reinforcement, these weaknesses can be mitigated, leading to more reliable and secure AI applications in the business domain.

It should be noted that our results cannot be generalized because it would require more experiments, a larger and more diverse sample of domain-specific questions, the possibility of assessment by a larger group of managers with different knowledge, and the inclusion of more dialogue assessment measures.

Conclusions

The goal of this paper was to overview of the current state of research and projects on intelligent interface of information systems. Particularly, the presentation was focused on Generative AI and its prospects to be used by managers. The state-of-the-art in Human-Computer Interactions was shown in the background of interface evolution from the first Decision Support Systems and Semantic Networks to the future Superhuman Manager. The conceptual framework of a Superintelligent Manager, conjoining the abilities of robots and humans, was sketched.

The most important part of the paper was concentrated on large language models used to understand queries, solve management problems and generate answers. Particular attention was given to two leading Generative AII platforms: ChatGPT and Bard. The methodology of usage of new communication facilities in management information systems was proposed and illustrated by dialog examples. The selected examples of interactions demonstrated high intelligence in dialog models and search engines. The crucial question concerned the evaluation of dialog quality. In our project, the subjective metrics of quality were proposed, namely: sensibleness, specificity, and interestingness measures. The examples demonstrated the advantages of the new technology in management and the weaknesses of current solutions. The preliminary evaluation of generated answers shows a big interest in large language models in management practice.

Today the research progress on dialog models and applications of Generative AI is extremely dynamic. Therefore, it is difficult to sketch a long-time perspective. However, our future research will be oriented towards a methodology of manager dialogs and semantic queries using domain knowledge and external data sources. In consequence, the corpus of queries and case studies will be extended, and more sophisticated evaluation measures will be designed, covering safety, groundedness, informativeness, and citation accuracy.

References

Adiwardana, D. et al. (2020). Towards a human-like open-domain chatbot. *arXiv:2001.09977*.
doi:10.48550/arXiv.2001.09977
Alexopoulos, P. (2020). *Semantic Modeling for Data. Avoiding Pitfalls and Breaking Dilemmas*. O'Reilly Media.
Babcock, J., & Bali, R. (2021). *Generative AI with Python and TensorFlow 2*. Packt Publishing.

Berners-Lee, T., Hendler, J., & Lassila, O. (2001). The Semantic Web. *Scientific American*, *284*(5), 34–43.

Bianchi, C. (2023). *Global Search Engine Desktop Market Share*. Retrieved from https://www.statista.com/statistics/216573/worldwide-market-share-of-search-engines/

Bollacker, K., Evans, C., Paritosh, P., Sturge, T., & Taylor, J. (2008). Freebase: A collaboratively created graph database for structuring human knowledge. *SIGMOD 08 Proceedings of the 2008 ACM SIGMOD International Conference on Management of Data*, 1247–1250.

Bostrom, N. (2016). *Superintelligence: Paths, Dangers, Strategies*. Oxford: Oxford University Press.

Brown, T. et al. (2020). Language models are few-shot learners. *Advances in Neural Information Processing Systems*, *33*, 1877–1901.

Canbek, N.G., & Mutlu, M.E. (2016). On the track of artificial intelligence: Learning with intelligent personal assistants. *Journal of Human Sciences*, *13*(1), 592–601. **doi:10.14687/ijhs.v13i1.3549**

Chien, C.F., Dauzère-Pérès, S., Huh, W.T., Jang, Y.J., & Morrison, J.R. (2020). Artificial intelligence in manufacturing and logistics systems: Algorithms, applications, and case studies. *International Journal of Production Research*, *58*(9), 2730–2731. **doi:10.1080/00207543.2020.1752488**

Conroy, S. (2023). *Google Bard and AI integration with Google Maps*. Retrieved from https://www.wepc.com/tips/google-bard-with-google-maps/

Cook, M. (2022). *Benchmarking: Diffbot Knowledge Graph Versus Google Knowledge Graph*. Retrieved from https://blog.diffbot.com/author/merrill/

Foster, D. (2019). *Generative Deep Learning: Teaching Machines to Paint, Write, Compose, and Play*. O'Reilly Media.

Fraley, A. (2023). *The Artificial Intelligence and Generative AI Bible*. Int. Kindle Ed.

Gopalakrishnan, K. et al. (2019). Topical-Chat: Towards Knowledge-Grounded Open-Domain Conversations. *Proc. Interspeech*, 1891–1895.

Gulliksen, J., Cajander, Å., Sad, B., Eriksson, E., & Kavathatzopoulos, I. (2009). User-centered systems design as organizational change: A longitudinal action research project to improve usability and the computerized work environment in a public authority. *International Journal of Technology and Human Interaction*, *5*(3), 13–53.

Hoffman, R. (2023). *Impromptu: Amplifying Our Humanity Through AI*. Dallepedia

Hussien, A., Rahma, A., & Wahab, A. (2021). Recommendation systems for e-commerce systems: An overview. *Journal of Physics. Conference Series*, *1897*. **doi:10.1088/1742-6596/1897/1/012024**

James, L. (2023). *8 Steps for an Effective Decision Making Process*. Retrieved from https://www.londonbusinessmag.co.uk/effective-decision-making

Kurzweil, R. (2006). *Singularity Is Near. When Humans Transcend Biology*. Springer.

Lopatovska, I. (2019). Overview of the Intelligent Personal Assistants. *Ukrainian Journal on Library and Information Science*. **doi:10.31866/2616-7654.3.2019.169669**

Min, B. et al. (2021). *Recent advances in natural language processing via large pre-trained language models: A survey*. **doi:10.48550/arXiv.2111.01243**

Pawełoszek, I., & Korczak, J. (2023). Future Manager – Perspective of Human and Artificial Intelligence, Futures, (to be published).

Piotrowski, D. (2022). Demographic and Socio-Economic Factors as Barriers to Robo-Advisory Acceptance in Poland. *Annales Universitatis Mariae Curie-Skłodowska, sectio H – Oeconomia*, 56(3), 109–126. doi:10.17951/h.2022.56.3.109-126

Popovic, M. (2023). *Automate Your Decision-Making with ChatGPT*. Retrieved from https://docs.kanaries.net/articles/chatgpt-auto-decision-making

Power, D.J. (2008). Decision support systems: A historical overview. In P. Bernus, J. Blazewicz, G. Schmidt & M.J. Shaw (Eds.), *Handbook on Decision Support Systems 1* (pp. 121–140). Berlin – Heidelberg: Springer. **doi:10.1007/978-3-540-48713-5_7**

Roehl, U.B.U. (2022). Understanding automated decision-making in the public sector: A classification of automated, administrative decision-making. In G. Juell-Skielse, I. Lindgren & M. Åkesson (Eds.), *Service Automation in the Public Sector. Progress in IS*. Springer. **doi:10.1007/978-3-030-92644-1_3**

Singhal, A. (2012). Introducing the Knowledge Graph: things, not strings. *The Keyword* (blog), Google. May 26. Retrieved from https://blog.google/products/search/introducing-knowledge-graph-things-not/

Sordoni, A. et al. (2015). A Neural Network Approach to Context-Sensitive Generation of Conversational Responses. *Proceedings of the Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, *NAACL HLT 2015* (pp. 196–205). 31 May – 5 June 2015. Denver: Association for Computational Linguistics (ACL).

Stohr, E.A., & White, N. (1983). User Interfaces for Decision Support Systems: An Overview. *NYU Working Paper No. IS-82-63*. Retrieved from https://ssrn.com/abstract=1290182

Tarus, J.K., Niu, Z., & Mustafa, G. (2018). Knowledge-based recommendation: A review of ontology-based recommender systems for e-learning. *Artificial Intelligence Review*, *50*, 21–48. **doi:10.1007/s10462-017-9539-5**

Thoppilan, R. et al. (2022). LaMDA: Language Models for Dialog Applications. *Computation and Language*. doi:10.48550/arXiv.2201.08239

Torrey, L., & Shavlik, J. (2010). Transfer Learning. In *Handbook of Research on Machine Learning Applications and Trends: Algorithms, Methods, and Techniques* (pp. 242–264). Hershey IGI Global. **doi:10.4018/978-1-60566-766-9.ch011**

Vinyals, O., & Le, Q.V. (2015). A neural conversational model. *arXiv:1506.05869*. **doi:10.48550/arXiv.1506.05869**

Wang, H. et al. (2022). Enabling Harmonious Human-Machine Interaction with Visual-Context Augmented Dialogue System: A Review. *arXiv*, *abs/2207.00782*. **doi:10.48550/arXiv.2207.00782**

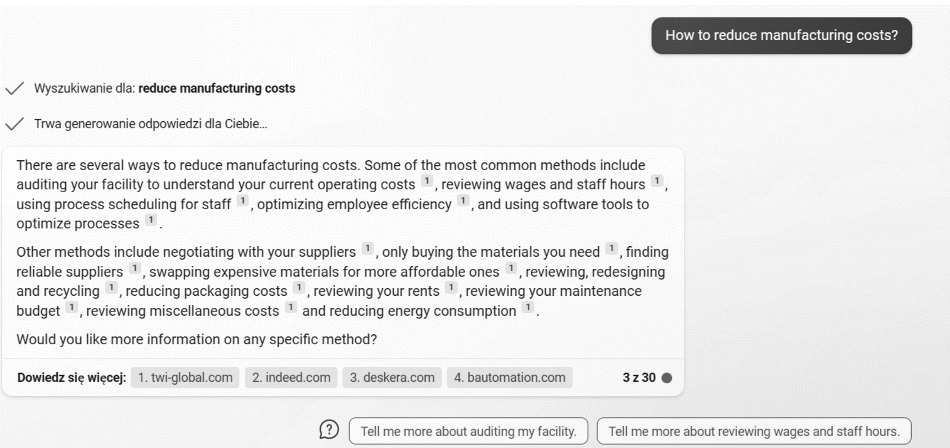
Wolfram, S. (2023). *What Is ChatGPT Doing ... and Why Does It Work?* Wolfram Media.

Zhao, W.X. et al. (2023). A Survey of Large Language Models. *arXiv*, *abs/2303.18223*.

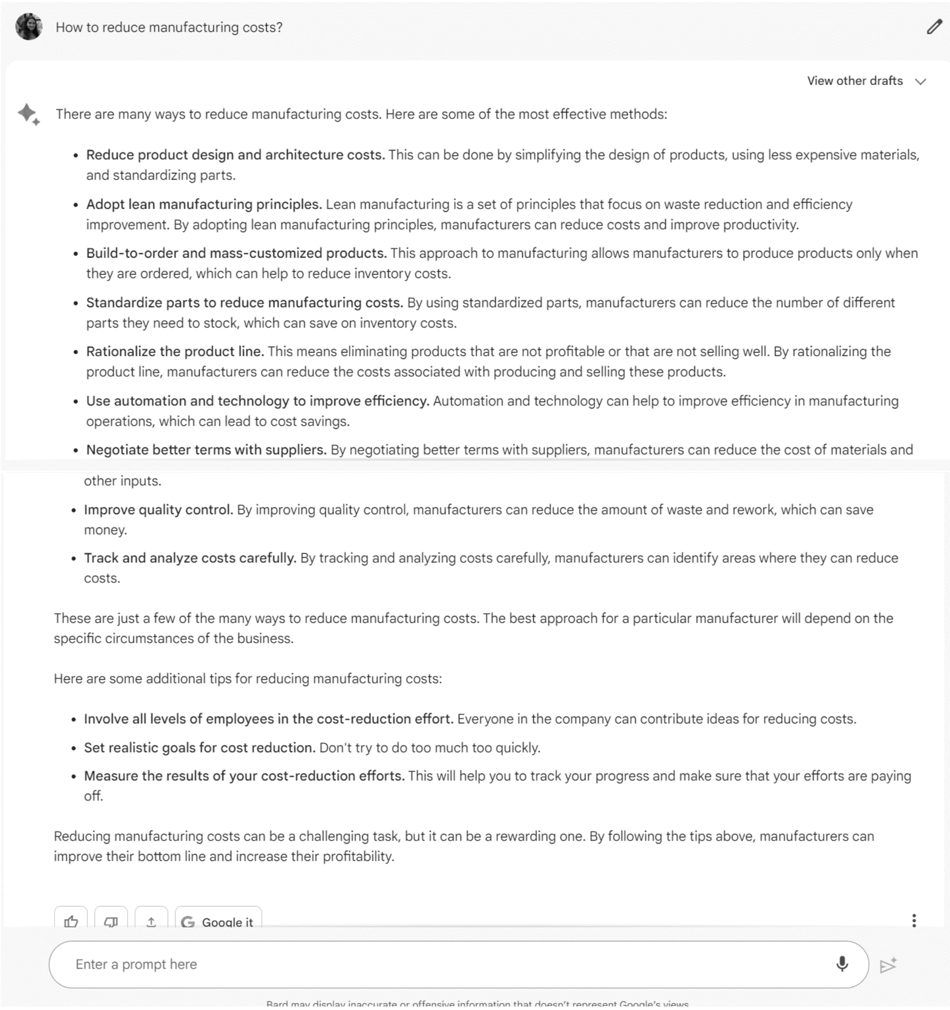
Zinczuk B. (2018). Artificial intelligence and its socially responsible use in the modern economy. *Annales Universitatis Mariae Curie-Skłodowska, sectio H – Oeconomia*, *52*(5), 125–133. doi:10.17951/h.2018.52.5.125-133

Appendix: Screenshots of selected dialog examples: A1, A2, A3

A1: Manufacturing costs – Chat GPT Bing



The answer is sensible, interesting, and well-grounded as the references are presented. The answer is focused mainly on operational issues.



Bard

Bard's answer is also sensible and informative, however, it lacks explicit grounding. The answers are more focused on strategic issues than on operational ones. Implementation of LLM as an internal decision support tool in an enterprise would require its integration with databases of ERP systems. Such a solution would make it possible to analyze costs and identify specific items that require changes.

A2: CEO candidates – Chat GPT

how to find the best candidate for the position of CEO?

Wyszukiwanie dla: **how to find the best candidate for the position of CEO**

Trwa generowanie odpowiedzi dla Ciebie...

There are several ways to find the best candidate for the position of CEO. One way is to choose a top-tier headhunting company to identify the best candidates [1]. It's important to define what the business needs from its next CEO and determine which stakeholders will play roles in the process and at which stages [1].

Board members who are adept at picking CEOs do four things others don't: They work painstakingly to clarify the essential qualities needed to succeed in the job; they keep an open mind about where the best candidate will come from; they go deep to understand which candidate is the best fit; and they allow for imperfections in the chosen candidate [2].

Is there anything else you would like to know?

Dowiedz się więcej: 1. mckinsey.com 2. hbr.org **1 z 30**

Chat GPT4 is short and informative and suggests further reading in trusted sources.

Bard

How to find the best candidate for the position of CEO?

View other drafts

Finding the best candidate for the position of CEO is a critical task for any organization. The CEO is responsible for the overall direction and success of the company, so it is important to find someone who has the skills, experience, and vision to lead the company to new heights.

Here are some tips on how to find the best candidate for the position of CEO:

1. **Start by defining the company's needs.** What are the most important qualities that the company is looking for in a CEO? What are the company's goals and objectives for the next few years? Once you have a clear understanding of the company's needs, you can start to identify potential candidates.
2. **Cast a wide net.** Don't just look for candidates who are already working within your industry. Consider candidates from other industries who have the skills and experience that you are looking for. You may also want to consider hiring a search firm to help you identify potential candidates.
3. **Conduct thorough interviews.** Once you have identified a few potential candidates, it is important to conduct thorough interviews. This will give you a chance to assess the candidates' skills, experience, and fit for the company culture. Be sure to ask questions that will help you to determine the candidates' leadership abilities, strategic thinking, and decision-making skills.
4. **Get input from others.** Once you have conducted the interviews, it is important to get input from others. This could include other members of the board of directors, senior executives, or even employees. Getting input from others will help you to get a more well-rounded view of the candidates.
5. **Make an informed decision.** Once you have gathered all of the information, it is time to make an informed decision. Be sure to weigh all of the factors carefully before making a final decision.

Finding the best candidate for the position of CEO is not an easy task, but it is an important one. By following these tips, you can increase your chances of finding the right person to lead your company to success.

Here are some additional tips for finding the best candidate for the position of CEO:

- **Look for candidates who have a strong track record of success.** This could include success in previous leadership roles, as well as success in achieving specific goals and objectives.
- **Look for candidates who are visionary and have a clear sense of direction.** The CEO should be able to articulate a clear vision for the company and develop a strategy to achieve that vision.
- **Look for candidates who are strong communicators and have the ability to motivate and inspire others.** The CEO should be able to communicate the company's vision to employees and stakeholders, and motivate them to achieve the company's goals.

- **Look for candidates who are ethical and have a strong moral compass.** The CEO should be someone who can be trusted to make decisions that are in the best interests of the company and its stakeholders.

Finding the best candidate for the position of CEO is a critical task, but it is one that can be accomplished by following these tips. By taking the time to find the right person, you can ensure that your company is well-positioned for success in the years to come.

Enter a prompt here

Bard's answer is more specific about the steps of the recruitment process and the requirements for the position of CEO. However, it does not point to the references and sources of knowledge. Here the connection with headhunting services and social media would be helpful.

A3: Grocery store – Chat GPT

How to choose the best location for autonomous grocery store in Częstochowa?

Wyszukiwanie dla: **how to choose the best location for autonomous grocery store in Częstochowa**

Wyszukiwanie dla: **how to choose the best location for a grocery store**

Trwa generowanie odpowiedzi dla Ciebie...

Choosing the best location for an autonomous grocery store in Częstochowa would involve considering several factors. You should study the number of people who live in the area, and how many of them can be your potential customers. You can research their probable incomes and whether they can afford your grocery products. Also, know about the average foot and vehicle traffic [1].

Your grocery store should be near the road so that people can walk or drive there easily. There should be ample parking space. Ideally, your shop should be located near a major road with good traffic [1].

Is there anything else you would like to know?

Dowiedz się więcej: 1. okcredit.in **2 z 30**

The answer is very general and could as well be applied to any other city. GPT4 lacks current knowledge about demography, traffic and urban planning.

Bart

How to choose the best location for autonomous grocery store in Częstochowa?

View other drafts

Here are some factors to consider when choosing the best location for an autonomous grocery store in Częstochowa:

- **Demographics:** The ideal location will be in an area with a high density of potential customers. This could include areas with a large population, a high median income, or a large number of young professionals.
- **Proximity to other businesses:** The location should be close to other businesses that attract customers, such as restaurants, gyms, or coffee shops. This will help to increase foot traffic and drive sales.
- **Accessibility:** The location should be easy to access by car, public transportation, or walking. This will make it convenient for customers to shop at the store.
- **Visibility:** The location should be visible from the street or from other high-traffic areas. This will help to attract customers and draw attention to the store.
- **Parking:** The location should have adequate parking for customers. This will make it easier for customers to shop at the store and will reduce the chances of customers driving away due to lack of parking.
- **Size:** The location should be large enough to accommodate the store's inventory and equipment. However, it should not be too large, as this could lead to higher operating costs.
- **Cost:** The location should be affordable. However, it is important to balance the cost of the location with the potential sales that the

store could generate.

Once you have considered these factors, you can start to narrow down your options and find the best location for your autonomous grocery store in Częstochowa.

Here are some additional tips for choosing the best location for an autonomous grocery store:

- **Do your research.** Before you start looking at locations, it is important to do your research and understand the market. This includes understanding the demographics of the area, the competition, and the potential sales that the store could generate.
- **Get feedback from others.** Once you have found a few potential locations, it is a good idea to get feedback from others. This could include potential customers, business owners, or even city officials. Getting feedback from others will help you to make an informed decision.
- **Be flexible.** It is important to be flexible when choosing a location. The ideal location may not be available, so you may need to be willing to compromise on some of the factors that you are looking for.

Choosing the best location for an autonomous grocery store is an important decision. By following these tips, you can increase your chances of finding the perfect location for your store.

Google it

Bard describes the main factors of the location. Later in the conversation, we asked to indicate specific places in Częstochowa where an autonomous store could be opened. Several suggestions were received, along with photos of places. However, it can be said that geography is not Bard's strong point, as the suggested locations were in other major cities. Enriching the language model with multi-criteria spatial analysis skills would require access to geographic information systems resources. Such initiatives are already underway. In particular, the planned integration of Bard with Google Maps deserves attention (Conroy, 2023).