

Exploring GPT-4's Reasoning Capabilities and Practical Applications for Business Integration

April. 20th 2023

Presented by:



Marc Taccolini

Founder and CTO @Tatsoft

"Exploring GPT-4's Reasoning Capabilities and Practical Guide for Business Adoption"

Moderator: Welcome everyone to today's webinar, "Exploring GPT-4's Reasoning Capabilities and Practical Applications for Business Integration." My name is Isabela and I'll be your moderator for today's session.

About Taccolini and Tatsoft



Marc Taccolini

Founder and CTO @Tatsoft
Founder of InduSoft (currently part by Aveva)
Three decades of experience in Industrial Automation Software
Academic background on Engineering and AI Logical Languages



FRAMEWORX

The Affordable Platform with Limitless Possibilities
for Enabling Industrial Applications

We're delighted to have Marc with us, who is the Founder of Tatsoft, an expert in real-time software for mission critical applications, with an academic background in AI Logical Languages, and daily experience working with GPT-4 and new technologies on its business processes and software products. Marc, thank you for joining us today.

Presenter: Thank you, Isabela. I'm excited to be here and share my insights on GPT-4's logical and reasoning abilities and how businesses can leverage its capabilities.

Moderator: Before we dive into the presentation, Marc, could you briefly explain what GPT-4 is and why it has become such a hot topic in the AI community?



Exploring GPT-4's Reasoning Capabilities and Practical Applications for Business Integration

TATSOFT
Powerful Industrial Software

GPT-4: Generative Pre-trained Transformer 4, with early signs of AIG

Human Life transformation innovations

Fire, Arc and bow, Wheel, Written Language, Gutenberg Press, Industrial Revolution, Computers, Internet, AIG



FACTORYSTUDIO

© 2023 TATSOFT, LLC

FRAMEWORX

Presenter: Absolutely. GPT-4, or Generative Pre-trained Transformer 4, is a state-of-the-art AI language model developed

by OpenAI. It's capable of understanding and generating human-like interactions, making it a powerful tool for various applications like content creation, translation, and much more. The concept of Natural Language Interaction with computer systems isn't new at all; even I, myself, three decades ago, was doing research on a related topic, at that time called Logical Languages. The groundbreaking event is that now, for the first time in history, this technology has reached the threshold to have sparks of AGI, which is an AI capable of having generalized human cognitive abilities, allowing natural language interactions as fluid as you would have with another person. Its potential to revolutionize the entire human society and certainly how businesses operate makes it a very hot topic in the AI community and the general public.

Moderator: That sounds fascinating. I'm sure our audience is eager to learn more about GPT-4's capabilities and how they can apply it to their businesses. Without further ado, let's begin the presentation. Please give us an overview of the topics you are covering.



Exploring GPT-4's Reasoning Capabilities and Practical Applications for Business Integration

TATSOFT
Powerful Industrial Software

Agenda

Part 1: Evaluating Logical Reasoning Capabilities in ChatGPT (3.5) and GPT4

- Formal Evaluation Test
- Results and Insights

Reference: Technical article at LinkedIn:

<https://www.linkedin.com/pulse/evaluating-logical-reasoning-capabilities-gpt-35-gpt-4-taccolini>

Part 2: Practical Guide for Business Adoption

- Typical Use Cases and Disruption
- Preparing your Organization
- Practical Considerations
- Wrap-up, session example

Q&A



Logical and Reasoning - Evaluation Test

Test Structure: 80 questions assessing different skills:

1. Vocabulary and semantic similarity
2. Pattern recognition using number sequences
3. Logical reasoning, specifically identifying word order in sentences
4. Pattern completion, specifically using groups and numbers.
5. Pattern recognition, identifying the next letter in a sequence
6. Analogical reasoning, specifically completing analogies.
7. Logical reasoning, specifically identifying the next word in a sequence or group.
8. Basic arithmetic, logical reasoning, and abstract thinking

Reference: Technical article at LinkedIn:

<https://www.linkedin.com/pulse/evaluating-logical-reasoning-capabilities-gpt-35-gpt-4-taccolini>

Presenter: First, let's examine its logical reasoning capabilities. I performed formal tests on both ChatGPT (3.5) and GPT-4 on its reasoning abilities, and I will share my conclusions and insights. Then, I'll give some simple, actionable management hints on how to start incorporating GPT-4 into businesses.

Moderator: I have a question: Why, before going into the business applications, did you decide to start with a formal evaluation of its logical and reasoning capabilities?

Presenter: There is already a lot being published about its text generating and analysis capabilities, as well as its use as a programming or marketing aiding tool. So instead of jumping to those topics, I will focus on some other practical aspects outside the mainstream discussion, but essential to leverage this technology in business processes and mission-critical applications. When you aren't dealing with fictional writing but business tools, the minimum basic requirement is the ability to interact with logic and consistency.

MODERATOR: That makes sense. About the Reasoning Capabilities article you just published, which the link is shared on the screen. Can you summarize what it is about?

Link to Article: <https://www.linkedin.com/pulse/evaluating-logical-reasoning-capabilities-gpt-35-gpt-4-taccolini/>



Exploring GPT-4's Reasoning Capabilities and Practical Applications for Business Integration

TATSOFT
Powerful Industrial Software

Test Results And Insights

GPT-3.5 score: 59%

GPT-4 score: 71%. GPT-4 significantly outperforms ChatGPT (GPT-3.5) GPT-4 score: 71%

Insights

GPT-4 crossed the average threshold for Senior Software Developers

Strengths: Natural language understanding and knowledge to concisely output and summarize contents is a multitude of subjects and scenarios is astonishing.

Issues: Output stills contains "hallucinations", non-factual added contents, misinterpretation of context, poor math skills, struggles to output aiming specific goals.

Presenter: I had many hours of interactions with ChatGPT, either playing with technology or working with it in the company. That allowed me to develop some initial intuition on its capabilities and some insights on the best way to interact, but before more extensive usage on products and business processes, I decided that a more formal approach was necessary.

Therefore, I designed and executed a structured testing, inspired by the classic Wason selection task, which is a logic puzzle used to study human reasoning. I modified it to better suit the GPT-4 evaluation, including sections on pattern recognition, analogies, basic arithmetic, and prediction, and conducted a series of tests

with both ChatGPT (3.5) and GPT-4. The detailed explanation of the test sessions and examples of interactions are published in the technical article.

Moderator: And what were your findings?



Exploring GPT-4's Reasoning Capabilities and Practical Applications for Business Integration

TATSOFT
Powerful Industrial Software

Test Results And Insights

GPT-3.5 score: 59%

GPT-4 score: 71%. GPT-4 significantly outperforms ChatGPT (GPT-3.5) GPT-4 score: 71%

Insights

GPT-4 crossed the average threshold for Senior Software Developers

Strengths: Natural language understanding and knowledge to concisely output and summarize contents is a multitude of subjects and scenarios is astonishing.

Issues: Output stills contains "hallucinations", non-factual added contents, misinterpretation of context, poor math skills, struggles to output aiming specific goals.

Presenter: Both models demonstrated an impressive ability to reason logically, with GPT-4 significantly outperforming ChatGPT. GPT-4's responses were more coherent and accurate, even when dealing with complex reasoning tasks. GPT-3.5 scored around 59%, and GPT-4 around 71%. To put those numbers in context, applying similar tests to thousands of candidates for Software Designer and Senior Programmer positions for over three decades, I found a pattern that typically the ability level working in those areas is above 70%. On that criteria, GPT-3.5 would be below the minimum score for those tasks, and GPT-4 just passed.

Moderator: That's remarkable. It certainly sounds like GPT-4's reasoning capabilities are a step above its predecessors. Now,

let's move on to the practical applications for business integration. How can businesses start leveraging GPT-4's capabilities?



Exploring GPT-4's Reasoning Capabilities and Practical Applications for Business Integration

TATSOFT
Powerful Industrial Software

Typical business cases

Content generation: GPT-4 can create high-quality, human-like content for marketing campaigns, blog posts, or even technical documentation.

Customer service: By integrating GPT-4 into chatbots or help desks, businesses can provide accurate, personalized, and efficient support to their customers.

Data analysis: GPT-4's natural language understanding can help businesses analyze and interpret large datasets, extracting valuable insights for decision-making.

Workflow automation: GPT-4 can streamline various tasks, such as email drafting, appointment scheduling, and document summarization, freeing up time for employees to focus on more complex tasks.

Product Embedding: The technology can also be embedded in all kinds of products and packaged solutions.

FACTORYSTUDIO

© 2023 TATSOFT, LLC

FRAMEWORKX

Presenter: There are several typical ways businesses can integrate GPT-4 into their operations. Some of the most common use cases include:

Content generation: GPT-4 can create high-quality, human-like content for marketing campaigns, blog posts, or even technical documentation.

Customer service: By integrating GPT-4 into chatbots or help desks, businesses can provide accurate, personalized, and efficient support to their customers.

Data analysis: GPT-4's natural language understanding can help businesses analyze and interpret large datasets, extracting valuable insights for decision-making.

Workflow automation: GPT-4 can streamline various tasks, such as email drafting, appointment scheduling, and document summarization, freeing up time for employees to focus on more complex tasks.

Product Embedding: The technology can also be embedded in all kinds of products and packaged solutions.

Exploring GPT-4's Reasoning Capabilities and Practical Applications for Business Integration

TATSOFT
Powerful Industrial Software

Typical business cases

- Content generation: GPT-4 can generate high-quality, human-like content for marketing campaigns, blog posts, or even technical documentation.
- Customer service: Integrating GPT-4 into chatbots or help desks can provide accurate, personalized, and efficient responses to customer inquiries.
- Data analysis: GPT-4's natural language understanding can help businesses analyze and interpret large datasets, extracting valuable insights for data-driven decision-making.
- Workflow automation: GPT-4 can streamline various tasks, such as email drafting, appointment scheduling, and document summarization.
- Product solution: Embedding GPT-4 into products can enhance user experiences and create new, innovative solutions.

Disruptive Completely new Applications

FACTORYSTUDIO © 2023 TATSOFT, LLC FRAMEWORX

But, looking at that list, it feels very pale, not really conveying the impact it will have in transforming not only business but human society. We don't have much experience incorporating that level of disruptive technology; you can argue and compare what we live now with the advent of computers, the Internet, or even written language. So it is necessary to proceed with awareness and strategic thinking. It's not only about creating a new business utility to optimize existing processes or localized product enhancements, but it is about understanding and incorporating

new concepts in almost all aspects of human interactions and business activities.

Moderator: It is certainly overwhelming. How can businesses start incorporating GPT-4 into their workflows and culture?



Exploring GPT-4's Reasoning Capabilities and Practical Applications for Business Integration

TATSOFT
Powerful Industrial Software

Preparing your Organization

- Evaluating organizational readiness
- Creating a strategic plan for implementation
- Ensuring a seamless transition
- Training & Onboarding
 - Employee training and onboarding
 - Addressing ethical concerns and biases
 - Maintaining a human-centric approach

- Evaluating organizational readiness
- Creating a strategic plan for implementation
- Ensuring a seamless transition
 - Training & Onboarding
 - Employee training and onboarding
 - Addressing ethical concerns and biases
 - Maintaining a human-centric approach

Marc: If you ask GPT-4 itself, it has a generic set of basic and traditional steps, like: Identify suitable use cases, Research and select the right tools, training and adoption, Monitor and evaluate

performance. While those generic concepts can be applied, they don't grasp nor address the magnitude of the potential disruption and expansion of opportunities. Before starting to go into tools selections for specific use cases, you must have a Human-Centric-Approach, ensuring a seamless transition.

The slide shows some of the initial considerations to put in place. Clearly, the main key aspect is educating yourself and educating your team. The AI tools are great, but how useful or not they are at this point in the technology depends on the quality of human resources driving its usage and interactions.

Additionally, you must start collecting practical actions and detailed information on implementation; otherwise, the beautiful high-level plan will certainly fail. This shows some practical considerations on adoption, including educating the team on various topics:



Exploring GPT-4's Reasoning Capabilities and Practical Applications for Business Integration

TATSOFT
Powerful Industrial Software

Practical considerations

- Managing accounts and organizing chat sessions
- Understanding issues: "hallucination," context loss, content replacement, incompleteness of initial replies, "bidirectional laziness", addition of misinformation, boilerplate contents (writing a lot, not saying much), and others.
- Absolutely necessary human review.
- Properly executing chat sessions, API usage
- Prompt Engineering.

- Managing accounts and organizing chat sessions
- Understanding issues: "hallucination," context loss, content replacement, incompleteness of initial replies, "bidirectional laziness", addition of misinformation, boilerplate contents (writing a lot, not saying much), and others.
- Absolute need for human review.
- Properly executing chat sessions, API usage, and Prompt Engineering.

Moderator: There is plenty of ongoing discussion about Prompt Engineering; can you elaborate on it?

Marc: Sure, I'll start with an example from our logical test. The first time we gave it to GPT-4, it got only 41 out of 80, not 57, its final result. Some inputs in the first pass weren't created with the necessary clarity and context for the proper use by the models. The techniques for crafting effective prompts, managing chat interactions, and planning API integrations are fascinating but a bit too vast for us to address in this session.

Exploring GPT-4's Reasoning Capabilities and Practical Applications for Business Integration

TATSOFT
Powerful Industrial Software

Find Balance on the Opposites

Understand the Limitations of the AI

Don't assume any limitations

Keep the Pace

FACTORYSTUDIO

© 2023 TATSOFT, LLC

FRAMEWORKX

The slide features a central blue balance scale. The left pan is labeled 'Understand the Limitations of the AI' and the right pan is labeled 'Don't assume any limitations'. Below the scale, on the left, is a blue silhouette of a person running, with the text 'Keep the Pace' next to it. The slide includes a title at the top left, the TATSOFT logo at the top right, and footer text at the bottom: 'FACTORYSTUDIO', '© 2023 TATSOFT, LLC', and 'FRAMEWORKX'.

Moderator: Those are valuable tips. Thank you, Marc. I'm sure our audience now has a better understanding of GPT-4's reasoning capabilities and practical applications. Before we wrap up, do you have any final thoughts or advice to share?

- On the left side, Understanding the limitations of the AI
- On the right side, don't assume any limitations to the AI

Presenter: Embracing AI technology like GPT-4 is a journey that requires striking a balance between understanding its current limitations and not capping its potential at any pre-determined level. Businesses need to be prepared for both challenges and rewards along the way by fostering a culture of continuous learning and adaptation. The most successful applications will be those with fine-tuning that balance, as well as staying informed and agile to remain competitive in a landscape expected to evolve constantly and quickly.

As a wrap-up example, I'd like to briefly share some interactions I had with GPT-4 earlier today, brainstorming on IoT solutions, using Tatsoft's FrameworkX and FactoryStudio.



The screenshot shows a user interface for 'Exploring GPT-4's Reasoning Capabilities and Practical Applications for Business Integration'. The user 'MA' asks for product and documentation details. The AI response provides a product name and a structured documentation organization.

Exploring GPT-4's Reasoning Capabilities and Practical Applications for Business Integration

TATSOFT
Powerful Industrial Software

MA My company s launching a software that allows to define the configuration of Edge data collectors publishing data to the cloud, and the creation of dashboards using that data, using natural language. Create the name and the documentation organization for that product.

Product Name: CloudEdge Linguist

Documentation Organization:

1. Introduction
 - 1.1. Overview of CloudEdge Linguist
 - 1.2. Key Features
 - 1.3. System Requirements

FACTORYSTUDIO © 2023 TATSOFT, LLC FRAMEWORKX



MA

Using your training data, verify if SmartIoT is available to use, and decide if it is a better name. Then select the best communication protocol to use in the solution, then estimate the number of programmers necessary to develop it.



As an AI language model, I cannot verify in real-time if the name "SmartIoT" is available for use. You would need to perform a trademark search and domain name search to ensure that the name is not already in use by another company or product.



Regarding the communication protocol for your solution, it depends on your specific requirements and the target use cases. However, some popular and widely-used communication protocols for IoT and edge devices include MQTT, CoAP, and HTTP/REST. MQTT is lightweight and suitable for low-power devices, CoAP is designed for constrained networks, and HTTP/REST is widely adopted and easy to integrate with web services.



MA

For the first release, we need something more simple. So remove some or all of the Advance Features, select only one protocol from your list, and list only 5 types of dashboard gadgets that can be useful the segment of manufacturing.

2 >



For a simplified first release of the software, we can focus on the essential features, a single communication protocol, and a limited number of dashboard gadgets tailored for the manufacturing segment.



1. Communication Protocol: MQTT is an excellent choice for manufacturing environments, as it is lightweight, efficient, and suitable for low-power devices and constrained networks. It also provides real-time communication, which is essential for monitoring manufacturing processes.
2. Dashboard Gadgets:
 - a. Production Line Monitoring: Visualize the real-time status of each production line, including operational, idle, and faulted states.
 - b. Overall Equipment Effectiveness (OEE): Display a gauge or bar chart showing the OEE



Exploring GPT-4's Reasoning Capabilities and Practical Applications for Business Integration

To estimate the number of programmers needed to develop the solution, we need to consider the project's complexity, time constraints, and your budget. Here's a rough estimate:

1. Project Manager - 1
2. Backend Developers (cloud, data processing, and API) - 2-4
3. Frontend Developers (user interface, dashboard, and visualization) - 2-4
4. Embedded/IoT Developers (Edge data collectors and communication protocol) - 2-3
5. QA/Test Engineers - 2-3

In total, you may need around 9-15 programmers to develop the solution, assuming a moderate level of complexity and a reasonable timeline. Keep in mind that these numbers can vary based on the specific needs of your project.



MA

Can Tatsoft FrameworkX accelerate or enhance the development of that application?



Tatsoft FactoryStudio, also known as FrameworkX, is a powerful industrial software platform for building automation systems, SCADA, HMI, and IoT applications. It can accelerate and enhance the development of your natural language-driven configuration and dashboard creation application in several ways:



1. **Integrated Development Environment (IDE):** FactoryStudio provides a comprehensive IDE with a rich set of tools, components, and features for designing, developing, and deploying automation applications. This can help you create the user interface, logic, and data processing for your application more efficiently than building everything from scratch.
2. **Native MQTT Support:** FactoryStudio offers built-in support for the MQTT protocol, making it easy to integrate Edge data collectors and configure MQTT parameters directly within the platform. You can leverage FactoryStudio's MQTT features to manage connections, topics, and data transmission without writing custom code.
3. **Real-time Data Management:** FactoryStudio is designed to handle real-time data processing, storage, and visualization. You can use its powerful data management features to process data from Edge data collectors, store it in various formats (such as SQL databases, CSV files, or in-memory storage), and display it in real-time dashboards.



4. **Dashboard Components:** FactoryStudio includes a wide range of pre-built dashboard components (such as charts, tables, gauges, and maps) that can be easily customized and integrated into your application. This can save you time and effort in designing and implementing your dashboards.
5. **Scripting and Extensibility:** FactoryStudio supports scripting languages like VB.NET, C#, and JavaScript for implementing custom logic, algorithms, or integrations. You can use these scripting capabilities to create the natural language processing component of your application and incorporate it into the FactoryStudio environment.
6. **Scalability and Performance:** FactoryStudio is designed to handle large-scale, high-performance applications. It can manage thousands of devices, process millions of data points, and support multi-user environments. This scalability can be particularly beneficial for IoT applications with a large number of Edge data collectors and users.

In summary, Tatsoft FactoryStudio (FrameworkX) can accelerate and enhance the development of your natural language-driven configuration and dashboard creation application by providing a robust, scalable platform with built-in support for MQTT, real-time data management, and rich dashboard components. By leveraging these features, you can focus on the unique aspects of your application and reduce the time and effort required for development.



What is your estimate in reduction of resources to create the application, and on the maintenance and evolution life cycle using Tatsoft?



Using Tatsoft FactoryStudio FrameworkX to create the CloudEdge Linguist Lite application can lead to significant reductions in resources required for development, maintenance, and evolution throughout the application's life cycle. The platform's built-in features and libraries can streamline the development process and improve long-term maintainability.



Estimations of resource reduction with Tatsoft FactoryStudio FrameworkX:

1. **Development time:** By leveraging the built-in components, protocol support, and data management capabilities of FactoryStudio FrameworkX, you could potentially reduce the development time by 30% to 50%. This is because you won't have to build these components from scratch and can focus on the unique aspects of CloudEdge Linguist Lite, such as natural language processing and dashboard creation.
2. **Maintenance:** Using FactoryStudio FrameworkX can result in a more stable and maintainable application since it follows industry best practices and comes with support from Tatsoft. The platform's modular design can make it easier to manage updates and bug fixes, reducing maintenance effort by approximately 20% to 40%.



Exploring GPT-4's Reasoning Capabilities and Practical Applications for Business Integration

TATSOFT
Powerful Industrial Software

3. Evolution: As the application evolves over time, FactoryStudio FrameworkX's scalable architecture and built-in tools can make it easier to add new features and functionalities, resulting in a reduction of 20% to 30% in resources required for evolving the application.

Please note that these estimations are based on general experience with software development projects and may vary depending on the specific requirements and challenges of the CloudEdge Linguist Lite application.

Additionally, using Tatsoft FactoryStudio FrameworkX can provide indirect benefits such as improved application performance, better security, and increased reliability. These advantages can contribute to a more satisfying user experience and a higher adoption rate for the CloudEdge Linguist Lite application.

FACTORYSTUDIO

© 2023 TATSOFT, LLC

FRAMEWORX



Exploring GPT-4's Reasoning Capabilities and Practical Applications for Business Integration

TATSOFT
Powerful Industrial Software

Q & A

Visit us at tatsoft.com

FACTORYSTUDIO

© 2023 TATSOFT, LLC

FRAMEWORX