ORIGINAL RESEARCH ARTICLE

Visualization for a new era: Impact and application of large language models and AIGC to traditional business models

Qianqian Yang[1], Ngai Cheong[2,*], Dejiang Wang[3], Shi Li[4], Oi Neng Lei[5]

[1] *School of Economics and Finance, South China University of Technology, Guangzhou 510000, China*

[2] *Faculty of Applied Sciences, Macao Polytechnic University, Macau, China*

[3] *Intelligent Cloud Business Group, Infrastructure Department, Beijing 100000, China*

[4] *Department of Economics and Finance, City University of Hong Kong, Hong Kong, China*

[5] *Macau Industry-University-Research Collaboration Association, Macau, China*

* Corresponding author: Ngai Cheong, ncheong@mpu.edu.mo

ABSTRACT

This paper focuses on the application and business value of large-scale language models, such as GPT and Ernie's model. These models combined with AIGC tools like stable diffusion generate images with fixed styles, character traits, and continuous plots using randomized story scripts. As a result, it enhances the operational efficiency between or within industries widely, and it fully demonstrate their business value. On the technical side, this paper describes in detail of building a pipeline to generate cue words required for stable diffusion, in which using large-scale language models and story scripts. Subsequently, the limitations of text-to-image are summarized by comparing the traditional method and language model, i.e. comparing characteristics from traditional book production and images generated using language model's cue words. This leads to a supervised multiround iterative LoRA modeling scheme that utilizes CLIP to achieve character IP fixation. To evaluate the impact of the application direction, we combine application scenarios and researches on application aspects regarding current AIGC industry structure, we found that the AIGC tool has several major aspects, mainly includes the aspects of basic big model, industry and scenario models, business and domain small models, AI infrastructure and AIGC supporting services. big model and AIGC techniques generate images with no specific rules and have less limitation. We call this 'visualization' in the new AI era. In this paper, we explore the possible impacts and economic values when changing from traditional domain to the new AI ear.

Keywords: large-scale language models; AIGC tools; image generation; operational efficiency; conversion of text into customized pictures; visualization in the new AI era; application scenarios; LoRA modeling scheme

1. Introduction

Human civilization as a whole has undergone tremendous change and development, and in the process, technological innovations and breakthroughs have propelled society forward, with the emergence of the Large Language Model. It has provided us with a powerful tool to reduce costs and increase efficiency in most industries. But other studies concluded that, for the 32.8% of occupations, while incorporating AI technologies can increase productivity and efficiency; it may also lead to job displacement and unemployment, the effects may be particularly pronounced and disruptive[1]. In short, the Big Language model has a wide coverage and applicability that affects the operating models of various

industries. According to research in respect of food testing domain, a knowledge graph serving as an external knowledge base to support retrieval in the large language model, in order to overcome the problem of machine hallucination[2]. In terms of education, the larger language models can be used to create educational content, improve student engagement and interaction, and personalize learning experiences[3]. For media field, creating audio content with encompassing storylines speech, music, and sound effects, guided by text instructions leverages LLMs[4]. In addition, personalized services also take advantages because large language models can expand the scope of personalized services[5]. This paper will develop through the technical aspect, from introduction of the large language modeling tool and application examples, to realize the promotion of application scenarios at the technical level.

2. Related work

2.1. Large language model

At present, many companies and institutions have launched large language models with various characteristics. The GPT3.5 and GPT4 models are behind the most famous ChatGPT application. The Ernie model, which is more friendly to the Chinese context launched by Baidu, China. The LARG models have strong abilities in the field of literary creation; understanding for, the context and semantics of long story scripts, summarizing the scripts as a whole story and handling the secondary creation. In this paper, we will further explore its capabilities using a large language model coupled with the AIGC technique[1].

2.2. Diffusion model

The emergence of AIGC is inseparable from the application of diffusion model in this field, which also produced many excellent texts to image large-scale model tools. For instance, Midjourney released through discord and the open-source text to image large-scale model stable diffusion[3]. Although both have the capabilities to convert the excellent texts to images, but there are significant differences in many aspects. We ultimately chose stable-diffusion to work with the big language model to generate images.

3. Large language model text parsing and authoring

Parsing and secondary authoring of text by using large models has been applied in many domains, the popularity of large models have promoted the continuous emergence of various intelligent applications regarding copywriting[6]. Leonardo AI is a concrete example of generating the appropriate resources from a text-generated graph. To upload an image in Leonardo AI, when you click on the "Upload Image" button on the left toolbar, you can either upload an image from a local drive, or select something you have previously generated, or even something from a community feed. The image will appear on the canvas, and you can start editing it by adding it to the canvas. Next you can select and move. Using the Selection tool, click on the element you want to move (including the generated frame and any new images generated) to move the element on the canvas, drag and drop the element to the desired location, and click on the Pan button to seamlessly navigate around the canvas. Use the Draw Mask or Erase buttons on the toolbar to effectively modify the image. Masking an area preserves some of the theme and style underneath, making it an ideal area for complementary colors. Erasing an area removes unwanted parts or prepares them for replacement. After making changes using these tools, enter the image and regenerate the areas within the Generate frame to prompt and generate new content, enter your prompt in the "Prompt" text field and click the "Generate" button to prompt, specifically using the full sentence inputs. A good prompt describes all the elements you want to include in the generated frame, or you can use the "Negative Prompt" button to tell Al what not to generate. After generating the image, you can use the arrow buttons underneath the generated frame to cycle through the different variants. Another successful business example is the looka site that generates logos with AI. The steps to use it are simple, and it is suitable for many people to get started quickly. Go to looka, click on the

"Get Started for Free" button and enter your company name and industry type. On the next page, enter your company slogan (optional) and identifying keywords, and select your favorite icon style. On the next page, select the desired branding kit (optional) and color scheme, and edit the fonts and icons if required. With the library of icons and fonts provided by looka, elements such as colors, fonts, icons, etc. can be freely adjusted to meet individual needs.

It can provide a more powerful and deiversified basic model to customers. Hence, the upgrade of the TI platform tool chain ensures the efficiency and stability of training-tasks operation, and provides task management, model optimization and deployment support simultaneously. Therefore, customers can easier realize the implementation of the model. Through the upgrade of intelligent applications, Tencent Cloud provides customers with A variety of plug-and-play products are available. Combining with AI large model technology with actual business, it provides with more efficient, natural and intelligent solutions to clients.

Using large models to analyze and create text is crucial for stepping in a more efficient and innovative future; in which it can bring new traffic entrances and broader business opportunities to enterprises[7]. AIGC, one of the large models which based on intelligent algorithms and deep learning technology, can quickly generate high-quality text, attract a large number of users to click and share, increase the exposure and traffic of the platform. These impacts are vital to increase business opportunities.

While utilizing the Big Language Model in terms of article parsing and authoring capabilities, we will experiment and better understand how is a story text, parsing and recreating through the Big Language Model. The protagonists of this story are a little fox explorer and his bear friend (Appendix A.1). There are no case of using a widely circulated story that results in a large model having used it as training data, ensuring the fairness of both comprehension of the whole story and summarization of experimental process.

We use ChatGPT and ERNIE model respectively to understand and summarize the story script (in Chinese environment), perform split-scene processing on the story. The split-scene processing prompt is described as follows.

"Understand the story and describe it in the form of a storyboard script, which is divided by changes in the scene and accompanied by appropriate character dialogue. Remember to give a background description of the storyboard before it starts."

Overall, it seems that both models are in a usable state, on the other hand, they need to be manually fine-tuned or optimized in a secondary way according to their characteristics.

After obtaining the descriptions of the story subplots, we analyze the background and characters of the subplots and summarize the words respectively. Background analysis requires the large language model to describe the scenery of the language background, and in this process, both the ERNIE model and the GPT3.5 model will give the depictions of with emotions, faces, and other words that are not related to the scenery. In During this process, we can filter the word set twice（The use of this filter is consistent with the application of the CLIP model through scoring as mentioned in **Figure 1** below）, for example, by using the prompt: "Pick out words related to the landscape description".

Figure 1. Benchmark image and its style classification score table.

To make the big language model gives a more accurate set, or by using the shared filter. One of the storyboards descriptions and the corresponding word set as shown below.

- Sunlight, trees, leaves, forest floor, fox, explorer, nose, awakening. (Before screening);
- Sunlight, trees, leaves, forest floor. (Post-screening).

Using the same method, we can also get the corresponding set of descriptors for different characteristic. Using the power of big language modeling, we it can replace the manual process and parse textual plots by hand; Secondarily, transforms and releases human resources. To the extent, we can generate descriptive keywords and use AIGC technology to generate images for "visualization".

4. Combined with AIGC technology to generate customized images

Using the parsed lexical set produced by the large language model, we can use it as a prompt accelerator for to generate generating images for the AIGC tool[8]. Thanks to the text parsing and authoring capabilities of the big language model, each image can be plot-wise continuously. However, due to the instability of text to image generation, these images are still not guaranteed to be consistent in their drawing style and some features.

We use a feature set judgment and LoRA multi-round iteration scheme to solve the feature instability problem. Specifically, it can be divided into two parts, the first part is to initially describe and generate a usable character image as a baseline character benchmark image and use the feature set that links us using the CLIP model for character feature judgment. For example, let's first utilize the prompt: "masterpiece, (water color style:1.2), best quality, forest, a cute small humanoid fox explorer, (solo: 1.3), fantasy, dreamlike, big detailed eyes" and this prompt is a special type of text that can be used directly in the stable Diffusion and MidJourney. And then the multiple images that are matching the prompt description are generated, from which one is selected as the baseline character benchmark image; Select a subset of the feature set, for example, the age subset of the feature set has the elements of "infant, child, adolescent, adult, middle-aged, elderly", the CLIP model is used to classify the benchmark image with a single feature, and the corresponding classifications shown in **Figure 1**.

Based on the above classifications, we can judge the generated images, in order to generate other images that are similar to the current benchmark images, we need to use adolescent as the descriptor[9]. We can try to further classify the base images according to the classification of its subsets, its subsets construct with multiple feature sets in advance.

Using the multiple classifications, many feature dimensions can be obtained. Using the feature words with the highest scores in each subset of results, we can get a quadratically optimized prompt. Then, using the optimized prompt, along with the CharTurner LoRA (https://civitai.com/models/7252/charturnerbeta-lora-experimental) and a few other fixed parameters, we can get a series of more similar image sets. These image sets can be used as our initial training image set. After all that, we can use the Kohya_ss (https://github.com/wizardAEI/script-to-book-story/blob/main/example.ipynb) framework to train the initial version of LoRA model. But nevertheless, with the initial version of the LoRA model, still, there are cases where some of the features are not generated stably or are not yet fixed.

In order to solve the above problem, we trained LoRA with loop iterations as the following Equation (1):

$$G_n = train(G_{n-1}; x_n, \{A\}, picture_0) \tag{1}$$

The Equation represents the most basic LoRA iterative training process, where G is the LoRA model of the current iteration, starting from 0 represents the initial function, x represents the feature vector of the current training, and A is a fixed set of basic training words, including LoRA trigger words, non-generic feature words that want to be changed subsequently, etc.

It is easy to conclude from the above equation that when n tends to infinity training infinity, thereafter, a final model H can be obtained whose generated image features are identical to the benchmark image features as shown in the following Equation (2). In real training, we can decide when to stop the LoRA iterative process based on the performance of the LoRA model generation results after iteration according to the actual demand.

$$H = \lim_{n \to \infty} train(G_{n-1}; x_n, \{A\}, picture_0) \tag{2}$$

We tracked the overall training process and illustrated the effect of training by comparing the sample differences between versions (https://github.com/wizardAEI/script-to-book-story/blob/main/example.ipynb). **Figure 2** shows that the performance of the LoRA model becomes more and more stable and converges to the character characteristics of the baseline images as the versions continue to be iterated.
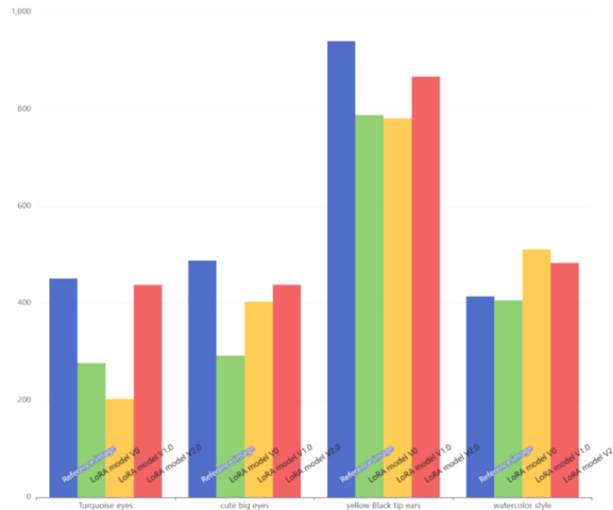

Figure 2. Compare the sample model generation of different iteration versions of LoRA with the original benchmark sample.

After obtaining a relatively usable version of LoRA, we use the model in conjunction with a collection of character description prompt that can be used to generate characters one by one. We can generate characters individually and then use character recognition and other solutions to select and overlay the task part into the background image that has already been generated according to the scene.

After obtaining a relatively usable version of LoRA, we use the model in conjunction with a collection of character description prompt[10]. we also can use the same background to produce some of the noise in order to carry out a partial repainting (inpainting 3), and to generate characters in a certain area. With plugins such as ContorlNet, we can better describe the character's pose (open-pose model) or expression (face model)[11]. Combining analyzing fairy tales and novels like the Ugly Duckling, Cinderella, etc. We can use a large language model to parse any literary work or article and then generate the corresponding continuous image, such as Appendix A2. The quality of the examples in the appendix is modelled using the CLIP model mentioned above: One specifies a score line until the model generation reaches that score line before it is adopted.

Techniques like this that which utilized large language models to parse text and generate uniformly styled multimodal media have been utilized in design software such as Canvas, Runway and Meitu[12]. Meitu's 2022 annual performance report is a strong proof of the technological developing development prospects of the technology. According to the financial report, the company's annual revenue was 2.085 billion yuan, an increase of 25.2% year-on-year, and the net profit was 111 million yuan, an increase of 29.9% year-on-year. Compared with the net loss of about 44.5 million yuan in the same period in 2021, for the first time, Meitu Company really turned a loss into a profit. It is worth mentioning that Meitu paid a dividend of about 78 million yuan for the first time, which reflects the confidence of Meitu's management in the sustainable and healthy

operation in the future, which can not only improve the expectations of the original investors for the company, but also help to further raise market funds and provide convenience for later refinance and additional shares. In the secondary capital market, Meitu's stock price has increased more than three times since its low in November last year, and its total market value has exceeded 11.4 billion Hong Kong dollars. The main reason why Meitu can achieve such results is that Meitu is a representative enterprise of many Chinese technology companies in the AIGC field. Since 2010, Meitu has continued to explore in related fields. Since the end of 2022, it has continuously achieved technical breakthroughs and carried out product implementation practice. With its technical advantages, Meitu can continuously expand the landing scene, open up leading imaging technology to individuals and small and medium-sized enterprises, and help the real economy realize digital transformation with the help of related businesses.

5. Conclusion

This big language model and AIGC technology solution turns the original large amount of text without rules and constraints into visual display pictures that can be generated by program scripting, which we can apply to many fields.

Using the Big Model's ability to understand and create text, coupled with AIGC's ability to text-generate diagrams, we can apply it in the traditional architectural design field to help clients visualize their building projects based on their descriptive text. Accelerate the design process. With the help of AI, a small outline can be provided based on a given content, and a complete copy can be quickly generated. Usually, more than 70% to 80% of the outline provided is available, so it can reduce the sorting of previous data and save time and human resources. At the same time, for architects, they may prefer to carry out the final finishing work. From the perspective of development, it will become more important to master this basic skill in the field of creative design and scheme design in the future. It can significantly improve efficiency faster and faster.

The small library science and technology team that is deeply cultivated in the construction industry are some enterprises that are deeply cultivated in the vertical field. We will take the construction industry as an example to discuss the landing path of AIGC in the vertical industry. At present, the domestic real economy is in the window period of transformation. "The integration of artificial intelligence and the real economy" is an important task[13]. Intelligent construction is based on the industrialization of new buildings, based on the deep integration of the new generation of information technology and advanced construction technology, running through all links of design, production, construction, operation and maintenance, supervision and other construction activities[14]. With the characteristics of self-perception, self-decision, self-implementation, self-adaptation, self-learning, etc., it can optimize the advanced construction methods of the quality, benefits and core competitiveness of the whole life cycle of the construction industry, and improve the development of the construction industry chain. The construction industry is a national pillar industry of nearly 30 trillion yuan per year, but its digitalization level ranks last among all industries in the country. Integrating intelligent buildings into it has the opportunity to reach a new level of "China Construction" and contribute greatly to China's economic development. Architectural design is an intellectually intensive service industry that is highly dependent on human resources, with manpower costs accounting for 70%–90% of the total. The arrival of the AI+ era is expected to improve industry efficiency and replace some of the mechanized drafting work. AI+ architectural design will significantly increase the per capita output value, reduce the size of the basic design staffing, reduce payroll expenses, and help reduce costs and increase efficiency, empowering the infrastructure design companies' Technology Under the wave of industry applications to accelerate the landing, according to the corresponding research report analysis, AI+ architectural design leading Huajian Group April 11 so far the maximum cumulative increase in share price of 129%. At present, in the construction industry is in a low state, Hua Jian Group's performance is steadily improving, "building+AI" driven business model,

medium- and long-term digital transformation is continued to be optimistic. China Design Group has also demonstrated the positive nature of "AI Architecture". The company released the interim report of 23 years, 23H1 achieved operating income of 2.239 billion yuan, −2.63% year-on-year, net profit of 262 million yuan, +1.36% year-on-year, net profit of 258 million yuan, +2.01% year-on-year, the profit growth rate is faster, we judge that it may be mainly benefited from the increase in the proportion of the high margin business, and the focus on the quality of the project brought about by the overall gross profit rate. Upward. In a single quarter, 23Q2 company achieved revenue/net profit of RMB 1.366/167 million, −1.17%/+1.15year on year The convertible bonds issued by the company were listed recently, and the construction of the Hoashi Innovation Centre project, which raised RMB 400 million, officially started on 9 June 2023, and is expected to further improve the company's technological strength after it is put into operation. The company in the field of intelligent city and intelligent transport continued to force, high gross profit business accounted for an increase in the proportion of "construction+AI" to create the second growth curve. 23H1 the company's traditional business of planning and research/survey and design/integrated testing/EPC and the whole process of the project management to achieve revenue of 2.8/12.4/2.4/1.2 billion yuan, respectively, −9.0% year-on-year, respectively, and −9.0%/2.2% year-on-year. Year-on-year respectively −9.0%/+2.3%/+9.2%/−18.9%, gross profit margin respectively 43.1%/40.5%/30.5%/19.1%, year-on-year respectively −5.0/+0.9/+8.6/−3.4 pct, low margin EPC business accounted for a further reduction. Among the emerging businesses, digital intelligence/low-carbon environmental protection business achieved revenue of RMB 130/190 million, +2.3%/−16.9% year-on-year respectively. Gross profit margin was 45.2%/20.1%, +9.3/+1.0 pct year-on-year[15]. **Figure 3** shows it.
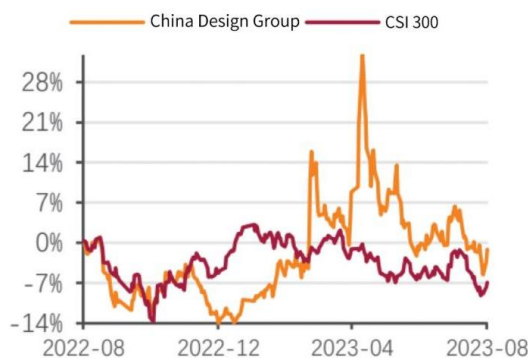

Figure 3. Performance of individual stocks against CSI 300 index.

And simulates medical scenarios to help medical students train. Or for simulation of medical diagnosis and treatment AI empowers diversified medical scenarios and drives the efficiency improvement of the medical industry chain. AIGC can also help achieve clinical path optimization, and find best practices and treatment patterns by analyzing a large amount of clinical data. This helps to reduce medical costs while improving the treatment effect of patients, and predicting treatment results based on patients' personal characteristics, disease types and treatment plans. Doctors can choose the most suitable clinical path and improve patient satisfaction. For doctors, AIGC is also a convenient medical guide library, which can help them keep abreast of the latest medical research progress, evidence-based medical evidences, and clinical guidelines, thus improving professional level and improving medical quality. To promote the development of industrialization and modernization, the emergence of ChatGPT is expected to accelerate the industrialization and commercialization of medical AI, and the value of AI in the vertical field will be fully recognized by the market. With the continuous implementation of AI technology in the medical field and the continuous promotion of the intelligent medical ecology, these enterprises can fly high and the stock price, promoting the circulation of cash flow and economic development as a whole[16]. The pharmaceutical sector rose against the trend, AI is conducive to the entire pharmaceutical industry to reduce costs and increase efficiency. With the

continuous implementation of artificial intelligence technology in the medical field and the continuous promotion of intelligent medical ecology, these enterprises can fly high, stock price, and enhancement. AI big model empowerment, medical+AI pioneer continues to lead. Shanghai Runda Medical Technology medical for AI first front runner, 2023H1 the company's R&D expense investment of 65,941,200 yuan, an increase of 11.85%, the company continued to increase the R&D investment in the field of IVD products and AI products, to meet the demand for clinical testing. 2023H1, the industrial plate achieved revenue of 289 million yuan, an increase of 36.63% year-on-year, the performance of the performance is bright. Among them, the company launched the "wisdom inspection - artificial intelligence interpretation test report system". AI products have been promoted and used in more than 40 medical institutions in China, and actively cooperate with Huawei and other domestic large model base company, to develop universal AI medical services large model, the relevant product development work is being steadily and smoothly promoted. In addition, it has also actively signed a formal cooperation agreement with Huawei Cloud, which is another landmark progress of the company in the field of AI and medical information technology, and continues to lead the industry. The company signed a comprehensive strategic cooperation agreement with Huawei Cloud, which will further empower traditional testing services to upgrade in the direction of refinement, informatization and intelligence based on information technology such as big data, cloud computing and artificial intelligence, and help hospital customers to further reduce costs and increase efficiency, as well as exploring the usage scenarios for C-users, accelerating the construction of Shanghai Runda Medical Technology's deep moat. Shanghai Runda Medical Technology Medical expects to achieve operating income of RMB12.341/14.682/17.274 billion in 2023–2025, + 17.6%/19.0%/17.7% year-on-year; net profit of RMB503/6.15/7.45 billion, +20.5%/22.2%/21.1% year-on-year; corresponding P/E 14/12/10x. /The valuation of this year's annual report maintains buy. Whether it is medical digitalization or pharmaceutical marketing empowerment, AIGC is leading us to a new era of enterprise digital marketing reform[17]. **Figure 4** shows it.

Figure 4. Company price comparison with SSE composite index.

Large Language Model plus AIGC likewise speeds up the scripting process and reduces costs for all kinds of movies, TV shows, and creative products. Many film and television leading companies have released the layout of AIGC; at the same time, film and television companies represented by Shanghai Film plan to strengthen IP development and operation. AIGC will be empowered to help the whole industry chain of content in the film and television industry. On the one hand, it is conducive to improving efficiency and reducing costs in the whole industry, and on the other hand, it is conducive to the deepening and amplifying of IP value. Earlier this year, SFG acquired 51% of the equity of ShangHai Film Infinite and gave them the 10-year operating rights of 30 classic IPs. Recently, SFG set up the "New Vision Digital Investment Fund for Cultural Industry", which intends to invest 300 million in three years for the incubation and landing of projects in the upstream and downstream of the cultural industry chain, with a special focus on the cutting-edge technologies

such as AIGC and Metacosmos. How to use technology to empower the story, so that the story's vitality continues, may be one of the directions of future film development.AIGC has entered the video era. In the future, it is expected to usher in a wide range of applications in many fields such as storyboarding, editing, special effects, cutscene animation, etc. and is expected to reshape the industry ecology together with text to image technology[18].

In summary, this research integrates and experiments with the use of big language models and AIGC technology as a 'visualization' tool for the new era. And it gives examples of how this combined form of technology can be used and generate value in various domains. We look forward to the development of this technology in the future and expect it to bring more convenience and to our lives. Providing digital inspiration to more industries, fuelling the development of the digital economy and fueling the development of the real economy. The future cannot be separated from the digital economy, the demand for economic and financial development will also become a driving factor in the development of big data models.

Author contributions

Conceptualization, DW, QY; methodology, QY, DW; software, DW; validation, QY, DW and SL; formal analysis, SL, DW, ONL and NC; investigation, QY and DW; resources, QY, SL and DW; data curation, DW; writing—original draft preparation, SL and DW; writing—review and editing, SL, DW, ONL and NC. All authors have read and agreed to the published version of the manuscript.

Acknowledgments

Conflict of interest

The authors declare no conflict of interest.

References

1. Zarifhonarvar A. Economics of ChatGPT: a labor market view on the occupational impact of artificial intelligence. Journal of Electronic Business & Digital Economics. Published online December 5, 2023. doi: 10.1108/jebde-10-2023-0021
2. Qi Z, Yu Y, Tu M, et al. FoodGPT: A Large Language Model in Food Testing Domain with Incremental Pre-training and Knowledge Graph Prompt. ArXiv. 2023, arXiv:2308.10173. doi: 10.48550/arXiv.2308.10173
3. Kasneci E, Sessler K, Küchemann S, et al. ChatGPT for good? On opportunities and challenges of large language models for education. Learning and Individual Differences. 2023; 103: 102274. doi: 10.1016/j.lindif.2023.102274
4. Liu X, Zhu Z, Liu H, et al. WavJourney: Compositional Audio Creation with Large Language Models. ArXiv. 2023, arXiv:2307.14335. doi: 10.48550/arXiv.2307.14335
5. Chen J, Liu Z, Huang X. When Large Language Models Meet Personalization: Perspectives of Challenges and Opportunities. ArXiv. 2023, arXiv:2307.16376. doi: 10.48550/arXiv.2307.16376
6. Li Y. Intelligent Environmental Art Design Combining Big Data and Artificial Intelligence. Complexity. 2021; 2021: 1-11. doi: 10.1155/2021/1606262
7. Törnberg P. How to use LLMs for Text Analysis. ArXiv. 2023, arXiv:2307.13106. doi: 10.48550/arXiv.2307.13106
8. Jin Z, Song Z. Generating coherent comic with rich story using ChatGPT and Stable Diffusion. ArXiv. 2023, arXiv:2305.11067. doi: 10.48550/arXiv.2307.11067
9. Hu EJ, Shen Y, Wallis P, et al. LoRA: Low-Rank Adaptation of Large Language Models. ArXiv. 2021, arXiv:2106.09685. doi: 10.48550/arXiv.2106.09685
10. Qi Z, Yu Y, Tu M, et al. FoodGPT: A Large Language Model in Food Testing Domain with Incremental Pre-training and Knowledge Graph Prompt. ArXiv. 2023, arXiv:2308.10173. doi: 10.48550/arXiv.2308.10173

11. Liu X, Zhu Z, Liu H, et al. WavJourney: Compositional Audio Creation with Large Language Models. ArXiv. 2023, arXiv:2307.14335. doi: 10.48550/arXiv.2307.14335

12. Lai T, Xie C, Ruan M, Wang Z, Lu H, Fu S. Influence of artificial intelligence in education on adolescents' social adaptability: The mediatory role of social support. Uddin MZ, ed. PLOS ONE. 2023; 18(3): e0283170. doi: 10.1371/journal.pone.0283170

13. Przybilla L, Klinker K, Lang M, Schreieck M, Wiesche M, Krcmar H. Design Thinking in Digital Innovation Projects—Exploring the Effects of Intangibility. IEEE Transactions on Engineering Management. 2022; 69(4): 1635-1649. doi: 10.1109/tem.2020.3036818

14. Zhang C, Zhang C, Zhang C, et al. A Complete Survey on Generative AI (AIGC): Is ChatGPT from GPT-4 to GPT-5 All You Need? ArXiv. 2023, arXiv:2303.11717. doi: 10.48550/arXiv.2303.11717

15. Available online: http://news.10jqka.com.cn/field/sr/20230830/44188511.shtml (accessed on 2 June 2023).

16. Lecler A, Duron L, Soyer P. Revolutionizing radiology with GPT-based models: Current applications, future possibilities and limitations of ChatGPT. Diagn Interv Imaging. 2023 Jun;104(6):269-274. doi: 10.1016/j.diii.2023.02.003

17. Available online: http://news.10jqka.com.cn/field/sr/20230607/43107125.shtml (accessed on 2 June 2023).

18. Zhang C, Zhang C, Li C. One Small Step for Generative AI, One Giant Leap for AGI: A Complete Survey on ChatGPT in AIGC Era. ArXiv. 2023, arXiv:2304.06488. doi: 10.48550/arXiv.2304.06488

Appendix

A.1 Scripts

Translated scripts:

Once upon a time, there was a little fox adventurer named Tiny. Tiny loved adventures and exploring new places. One day, Tiny decided to venture into the forest. In the forest, Tiny encountered an adorable little bear named Buzzy. Buzzy was a very friendly bear who greeted Tiny and invited them to go treasure hunting together. Excitedly, Tiny agreed. They journeyed through dense forests, passed through thickets and rivers, searching for treasures deep within the woods.

During their adventure, Tiny and Buzzy encountered many interesting things. They saw beautiful flowers and shimmering seashells in the river. Fearlessly, they pressed on, continuously searching for clues to the treasure. Finally, after much searching, Tiny and Buzzy found a sparkling treasure. The treasure was filled with beautiful gemstones and gold. The little bear and fox jumped with joy, having successfully completed their adventure.

A.2 Story understanding and storyboard generation by large language model and final images