



NUS
National University
of Singapore

EAI

EAST ASIAN INSTITUTE
NATIONAL UNIVERSITY OF SINGAPORE

EAI COMMENTARY

No. 78 2 May 2024

Sora's Surge: China's AI Race Heats Up in Response to OpenAI's Latest Breakthrough

By SHAN Wei

Synopsis

OpenAI's unveiling of Sora, a groundbreaking text-to-video AI model, triggers intense discussions on China's AI prowess. With its remarkable ability to generate intricate, emotionally nuanced videos from natural language descriptions, Sora sets a new benchmark, spurring Chinese tech giants like Baidu and ByteDance to intensify their AI race. Despite China's strides in AI research and development, challenges loom large, from computing power shortages to talent retention struggles and a lack of conducive environments for original breakthroughs. As China's government prioritises AI in its strategic plans, the quest to bridge the gap with Sora underscores the nation's ambition to shape the future of AI and propel innovation in the global landscape.

On 21 February 2024, a few days after the American company OpenAI launched Sora, its text-to-video Artificial Intelligence (AI) model, the trending topic "Has Sora Once Again Expanded The Gap Between Chinese And American AI?" surfaced on Weibo, China's largest Twitter-style social media platform.¹ A large number of netizens and

¹ <https://s.weibo.com/weibo?q=%23Sora%E5%8F%88%E6%8A%8A%E4%B8%AD%E7%BE%8EAI%E7%9A%84%E5%B7%AE%E8%B7%9D%E6%8B%89%E5%A4%A7%E4%BA%86%E5%90%97%23>, accessed March 2024.

many internet influencers engaged in heated debates over the potential impact of this new AI product on China.

Sora is another groundbreaking large AI model launched by OpenAI (following ChatGPT) that has shocked the world. Users simply need to describe scenarios in natural language and Sora can generate highly smooth, multi-perspective videos. These videos go beyond previous similar products with their intricately detailed backgrounds, elaborate multi-angle shots and emotionally nuanced characters.²

If ChatGPT has impressed the world by demonstrating remarkable understanding of human language, then Sora is astonishing for its understanding of the real world.³ OpenAI declares its ambition on its official website: Sora aims to be a simulator for the entire physical world. It has been trained to understand and simulate the physical laws in the real world “with the goal of training models that help people solve problems requiring real-world interaction”.⁴ Zhou Hongyi, founder of the Chinese cybersecurity company 360 Group, believed that Sora could potentially shorten the realisation of Artificial General Intelligence (AGI) from 10 years to one or two years.⁵ Elon Musk commented on social media, “G humans” (good game humans, or game over).⁶

In the few weeks since its inception, Sora has sparked widespread enthusiasm among investors, practitioners from various industries and on social media platforms in China. There have been heated discussions about how Sora is poised to reshape numerous industries, including film, television, gaming, video production, education, advertising and more. During the National People’s Congress (NPC) and Chinese People’s Political Consultative Conference (CPPCC) meetings (generally called the “Two Sessions”) held in March, Sora and artificial intelligence also became a hot topic of discussion among attending representatives.⁷

Chinese tech giants are closely watching Sora. Over a year ago, the emergence of ChatGPT had prompted Chinese companies to engage in a heated race to catch up. Within just a year, hundreds of large language models have been developed in China, but there still remains a significant gap with ChatGPT. The launch of Sora has once again brought a new wave of anxiety to Chinese tech companies, making them realise again they are being left behind by OpenAI.⁸ Amidst the technological rivalry between China and the United States, the competition surrounding Sora is poised to emerge as a new focal point between the two nations. Catching up with Sora has become a new track of AI race for Chinese tech companies.

² <https://openai.com/sora>, accessed March 2024.

³ <https://www.zhihu.com/question/645156038/answer/3420893602>, accessed March 2024.

⁴ <https://openai.com/sora>, accessed March 2024.

⁵ https://www.thepaper.cn/newsDetail_forward_26369095, accessed March 2024.

⁶ <https://twitter.com/elonmusk/status/1758375648605745411>, accessed March 2024.

⁷ <https://m.21jingji.com/article/20240226/herald/a9a8ec966c861809553418a8608199e8.html>, accessed March 2024.

⁸ <https://www.scmp.com/tech/big-tech/article/3253034/openais-sora-pours-cold-water-chinas-ai-dreams-text-video-advancements-prompt-more-soul-searching>, accessed March 2024.

MAJOR CHINESE PLAYERS IN THE RACE

Chinese tech companies are not entirely unprepared in the field of text-to-video AI. The most likely contenders to develop a Chinese version of Sora are companies like Baidu and ByteDance, which lead in the field of AI large models. Baidu is one of the earliest Chinese companies to establish a presence in the field of artificial intelligence. In 2019, when OpenAI released the second-generation GPT model, Baidu also introduced its ERNIE series of AI large models. In March 2023, Baidu launched its competitor to ChatGPT, the ERNIE bot. Although ERNIE bot initially focused on natural language understanding, it quickly incorporated multimodal capabilities, such as the ability to generate images and videos. In January 2024, Baidu introduced a new video generation model called UniVG, which can generate videos from text or images.⁹ However, these video models still cannot match Sora in handling complex scenes and depicting details. Videos generated by Baidu's AI often resemble pre-existing content retrieved from a library and stitched together to match textual content, rather than generating entirely new video content.¹⁰

In early 2024, ByteDance, the parent company of TikTok, released a high-definition text-to-video model called MagicVideo-V2. It is claimed that this model's clarity, coherence and understanding of textual semantics surpass mainstream text-to-video models currently available.¹¹ ByteDance boasts a monthly active user base of 743 million on Douyin¹² and over 1.092 billion on TikTok.¹³ The massive amount of short videos provides ByteDance with a huge database for developing its own video generation models. However, after the launch of Sora, ByteDance admitted that its in-house video generating AIs are still in a very early stage and lag behind leading counterparts by a significant margin.¹⁴

ChatGPT and Sora have reportedly put tremendous pressure on ByteDance's leadership. Zhang Yiming, the founder of ByteDance, has spent a considerable amount of time studying artificial intelligence. Faced with the potential threat of Sora in the video domain, Zhang and other senior management have emphasised to the entire company that this is a war that cannot be lost, urging the mobilisation of all resources to catch up with Sora.¹⁵ A few days after Sora's release, the media reported that Jiang Lu, a research scientist leading Google's video generation team, had joined ByteDance. Jiang is a key contributor in content generation for several services at Google and is considered one of the leading AI experts in the Chinese community.¹⁶

Other internet giants have also made their moves. In March 2024, Alibaba launched its own video generator, EMO, which imitated a video generated by Sora and had one of the characters

⁹ https://www.sohu.com/a/761274703_472308, accessed March 2024.

¹⁰ <https://mp.weixin.qq.com/s/ujlrq2Dp6gRwL7T9ihkXUQ>, accessed March 2024.

¹¹ https://www.sohu.com/a/761274703_472308, accessed March 2024.

¹² <https://www.questmobile.com.cn/research/report/1726888249161519105>, accessed March 2024.

¹³ <https://www.amz123.com/t/e8BocQ51>, accessed March 2024.

¹⁴ <https://www.scmp.com/tech/tech-trends/article/3253692/ai-takes-priority-tiktok-owner-bytedance-sora-disrupts-future-video-creation>, accessed March 2024.

¹⁵ <https://www.scmp.com/tech/tech-trends/article/3253692/ai-takes-priority-tiktok-owner-bytedance-sora-disrupts-future-video-creation>, accessed March 2024.

¹⁶ <https://www.scmp.com/tech/big-tech/article/3253137/former-google-video-generation-expert-joins-tiktok-owner-bytedance-amid-generative-ai-race>, accessed March 2024.

EAI COMMENTARY

singing.¹⁷ Tencent also introduced its video generation model, VideoCrafter2, in early 2024, which is capable of quickly creating high-definition videos of a few seconds based on user text input.¹⁸ Additionally, several startups have secured billions in investments. For example, founded in March 2023, Shengshu Tech focuses on multimodal large models, with a focus on text-to-image and image-to-video. In March 2024, the company received a new round of investments worth hundreds of millions, announcing that its large models “will reach the level of Sora within the year”.¹⁹

BEIJING’S AI AMBITION

Sora has attracted close attention from the Chinese government. During the Two Sessions held in March 2024, Sora and artificial intelligence became a hot topic among NPC deputies and CPPCC members. They discussed various aspects of how to catch up with Sora and make China an AI powerhouse.²⁰ For example, the chairmen of Xiaomi and iFlytek both suggested introducing AI courses to primary and secondary school education to accelerate the cultivation of AI talents. Some representatives also called for the establishment of open platforms for computing power and data sharing. Additionally, there were suggestions to increase the application of AI large models in various industries, such as promoting the integration of large models with manufacturing, energy, transportation, health care and other sectors.²¹

AI holds a priority position in the strategic plans of the Chinese government. In the ambitious “Made in China 2025” plan issued in 2015, the government aims to drive the manufacturing industry with AI technology, including robotics, information technology, new energy and aerospace industry, transitioning towards high-end manufacturing.²² In 2017, the State Council issued the “Next Generation Artificial Intelligence Development Plan” to position China as a leading global innovation centre in AI by 2030.²³ President Xi Jinping emphasised the importance of AI for China during a political bureau study session in 2018, stating that AI is a key driver in transforming China from high-speed development to high-quality development.²⁴

In September 2023, President Xi proposed the concept of “new quality productive forces” which became a new ideological buzzword during the Two Sessions in 2024.²⁵ The concept emphasises driving economic growth through innovation in science and technology,

¹⁷ <https://sea.mashable.com/tech/31479/alibabas-ai-video-generator-just-dunked-on-sora-by-making-the-sora-lady-sing>, accessed March 2024.

¹⁸ https://www.sohu.com/a/761274703_472308, accessed March 2024.

¹⁹ <https://mp.weixin.qq.com/s/edgNGnA6dW1jxcWlnoXJkA>, accessed March 2024.

²⁰ <https://www.scmp.com/tech/policy/article/3254851/two-sessions-2024-chinas-lawmakers-call-more-ai-development-catch-us-while-keeping-it-under>, accessed March 2024.

²¹ <https://wap.eastmoney.com/a/202403083005026446.html>; https://www.sznews.com/news/content/2024-03/07/content_30788629.htm; <https://m.21jingji.com/article/20240226/herald/a9a8ec966c861809553418a8608199e8.html>, accessed March 2024.

²² <https://www.abc.net.au/chinese/2018-10-08/xi-jinpings-plan-to-turn-china-into-the-ai-world-leader/10351674>, accessed March 2024.

²³ http://www.gov.cn/zhengce/content/2017-07/20/content_5211996.htm, accessed March 2024.

²⁴ http://www.xinhuanet.com/politics/2018-10/31/c_1123643321.htm, accessed March 2024.

²⁵ <https://www.scmp.com/tech/policy/article/3254080/chinas-new-productive-forces-concept-offers-fresh-hope-and-impetus-economic-transformation>, accessed March 2024.

EAI COMMENTARY

developing high-tech, high-efficiency and high-quality emerging industries and future industries. This concept demonstrates the efforts of Chinese leaders in their exploration of new paths for economic growth in China. In the context of the new quality productive forces, AI plays an important role. On 9 March, in response to the launch of Sora and the Two Sessions, China's state-owned Central Television Station interviewed Baidu CEO Li Yanhong. Li pointed out that AI is the core component of the new quality productive forces. He explained, "The so-called new quality productive forces refers to productivity and growth driven by innovation. The area with the greatest innovation potential is artificial intelligence technology".²⁶

At the NPC session, Premier Li Qiang proposed "AI+" in the government report, further elevating the status of artificial intelligence in China's development strategy. The concept of AI+ aims to accelerate the integration of artificial intelligence into various industries, becoming the infrastructure and core driving force for economic transformation and upgrading.²⁷ The "Made in China 2025" initiative advocates for the integration of AI into manufacturing, where manufacturing is at the centre and AI serves a supporting role to enhance efficiency and quality. Conversely, AI+ positions artificial intelligence at the centre, serving as the source of original innovation and economic growth, fostering new technologies, products and industries.²⁸ AI becomes the core engine driving new quality productive forces.²⁹

Over the past decade, China has experienced rapid growth in the field of AI. It was estimated that in 2023, the market size of AI industry in China would reach a scale of US\$25.14 billion, with a growth rate of 10.11%.³⁰ A large number of AI enterprises have emerged, totalling over 4,300. Significant achievements include the development of intelligent chips, frameworks and large-scale models.³¹

China has made enormous investments and experienced rapid growth in the research and development of AI. Overall, China tops the world in terms of the number of AI research papers and patent applications. According to the Stanford AI Index Report published in 2023, China has been the top producer of AI papers in the past five years, contributing 39.78% of AI papers published in journals worldwide in 2021, as well as 26.15% of conference papers in the same year. The European Union (plus the UK) and the United States ranked second and third in both categories.³² Additionally, the report states that China accounted for 51.69% of global AI patent applications in 2021, while the United States and Europe together accounted for only 20.81% of AI patents.³³

²⁶ <https://finance.eastmoney.com/a/202403103007145665.html>, accessed March 2024.

²⁷ <https://www.globaltimes.cn/page/202403/1308145.shtml>, accessed March 2024.

²⁸ <http://www.xinhuanet.com/20240308/7fcb3fe87b19412bae6df80a8f8cecf3/c.html>, accessed March 2024.

²⁹ <https://m.ofweek.com/ai/2024-03/ART-201700-8440-30628036.html>, accessed March 2024.

³⁰ As a reference, in 2023, the AI market size in the United States was about US\$37.23 billion, with a growth rate of 7.38%. <https://www.statista.com/outlook/tmo/artificial-intelligence/united-states#market-size> and <https://www.statista.com/outlook/tmo/artificial-intelligence/china#market-size>, accessed April 2024.

³¹ <https://m.ofweek.com/ai/2024-03/ART-201700-8440-30628036.html>, accessed March 2024.

³² https://aiindex.stanford.edu/wp-content/uploads/2023/04/HAI_AI-Index-Report_2023.pdf, accessed March 2024.

³³ https://aiindex.stanford.edu/wp-content/uploads/2023/04/HAI_AI-Index-Report_2023.pdf, accessed March 2024.

CHALLENGES

However, China's AI ambitions still face several challenges. The first is the limitation of computing power. Large-scale AI models require powerful computing support. The training of ChatGPT utilised at least 10,000 Nvidia A100 graphics processing units (GPUs).³⁴ The computing power used for Sora training has not been disclosed, but it is likely greater than that of ChatGPT. Several major US tech giants have purchased large quantities of Nvidia's most advanced GPUs, including A100, H100 and H200. These chips are among the products sanctioned by the United States against China, making it inaccessible for Chinese tech companies. According to Zhou Hongyi, China currently has approximately 500,000 GPUs nationwide, but companies like Meta or Microsoft in the United States may purchase millions of chips from Nvidia each, meaning that the computing power of the entire country of China is not comparable to that of a single US tech giant.³⁵ While China is still able to obtain the advanced GPUs from third markets like India, Taiwan and Singapore, the quantity is unable to meet the needs of training large AI systems.³⁶

The second challenge lies in the limitation of training data. China boasts the largest number of internet users globally and hosts some of the world's largest search engines, social networking sites and e-commerce platforms, generating massive amounts of big data. This serves as a crucial foundation for AI development. However, the availability of high-quality data for AI training in China is remarkably limited. This is partly due to reluctance among Chinese tech companies to share their data, resulting in a comparatively small number of public databases. Yet, a more fundamental reason stems from extensive censorship in China, encompassing stringent scrutiny over internet content, academic research, film production and artistic creation. This censorship restricts any content that the government disapproves of, while also targeting any material deemed sensitive by the authorities, even in non-political areas such as economics, society and culture.

Consequently, content on the Chinese internet is heavily biased, lacking diverse perspectives and viewpoints.³⁷ In-depth discussions in areas like history, culture and society often struggle to survive. This results in a shortage of high-quality Chinese language data.³⁸ This issue surfaced during the large language model race sparked by ChatGPT in 2023. Users found that many Chinese models have better recognition capacities for English than for Chinese. One team involved in the race candidly admitted that this was because the amount of Chinese data available for their training set was only about one-tenth of the amount of English data.³⁹ Training AI models for text-to-video generation requires high-quality materials such as images and videos. However, the related industries of video, film and television are tightly regulated in China. Therefore, Chinese competitors to Sora may encounter difficulties in obtaining high-quality image data that accurately reflects the reality of China.

³⁴ https://mp.weixin.qq.com/s/Dch_LFKNpl8lXg9pO_BZsQ, accessed March 2024.

³⁵ <https://m.cls.cn/detail/1602831>, accessed March 2024.

³⁶ <https://www.reuters.com/technology/chinas-military-government-acquire-nvidia-chips-despite-us-ban-2024-01-14/>, accessed April 2024.

³⁷ <https://www.aljazeera.com/economy/2023/3/2/china-wants-to-copy-chatgpts-success-censorship-makes-it-tricky>, accessed March 2024.

³⁸ Yang, E, 2023. The Digital Dictator's Dilemma. *Working Paper, Univ. Calif. San Diego*. https://www.eddieyang.net/research/AI_dilemma.pdf.

³⁹ <https://www.8btc.com/article/6810624>, accessed March 2024.

EAI COMMENTARY

The third major obstacle faced by the Chinese tech sector is the shortage of AI talent. The competition in technology may ultimately depend on the competition for talent. China's higher education system has rapidly shifted its focus to nurturing AI talent, with 34.4% (or 438) of Chinese universities offering AI majors as of 2022.⁴⁰ However, according to an AI talent white paper released in the same year,⁴¹ China still faces a shortage of five million professionals in the AI industry.

For China, retaining talent may be more important than nurturing talent. According to a report released by the Macro Polo Institute in 2023,⁴² the United States still maintains a significant lead in the number of top AI talents. In 2022, 42% of the top-tier AI researchers were working in companies and universities in the United States, compared to 28% in China (as shown in Figure 1). One fact that may unsettle Beijing is that nearly half of these top talents (47%) received their undergraduate degrees from Chinese colleges, while only 18% from US colleges.



A considerable portion of Chinese students choose to pursue graduate studies in the United States and remain abroad for work, which has made significant contributions to the talent pool in the United States. Among the top AI talents working in the United States, 38% were originally from China (37% were from the United States, as shown in Figure 2). In the team of 89 people behind ChatGPT, at least nine are from China.⁴³ In the 13-person team behind Sora, approximately four are from China.⁴⁴ Therefore, the main challenge facing China now is how to retain its talents to work in the country and attract talents from other countries. The Chinese

⁴⁰ https://www.sohu.com/a/537514082_104992, accessed March 2024.

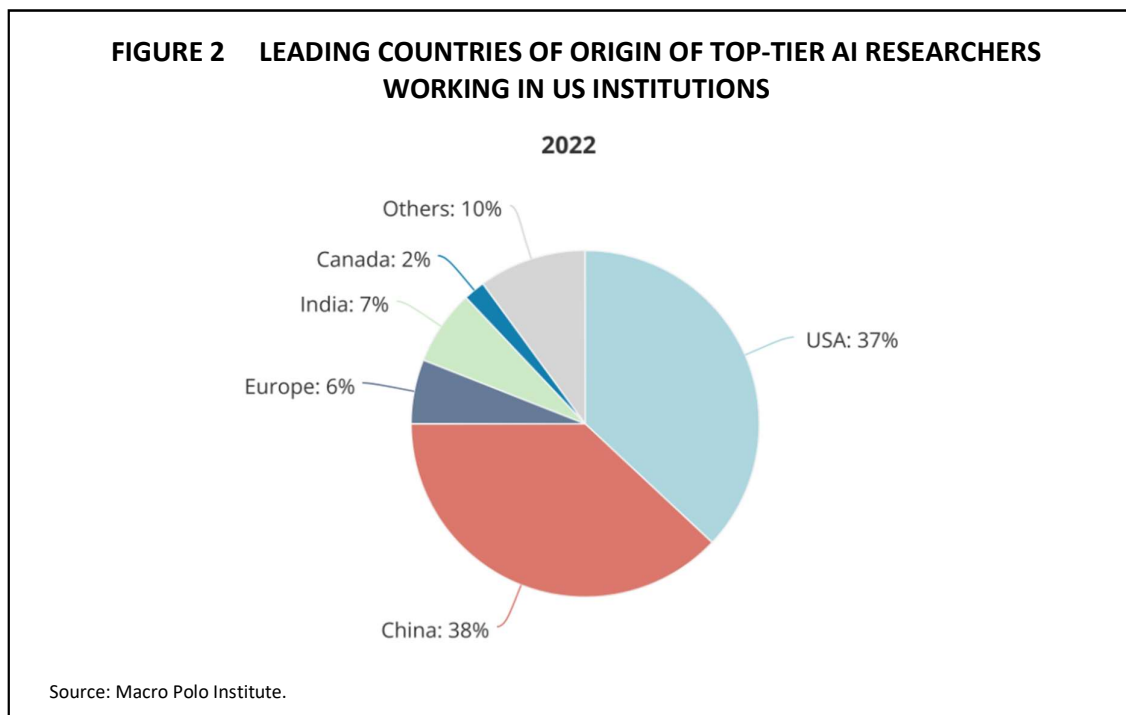
⁴¹ https://www.sohu.com/a/537514082_104992, accessed March 2024.

⁴² <https://macropolo.org/digital-projects/the-global-ai-talent-tracker/>, accessed March 2024.

⁴³ http://www.sohu.com/a/648578449_404981, accessed March 2024.

⁴⁴ <https://mp.weixin.qq.com/s/S1tuBaIIRGLZavym7plvpw>, accessed March 2024.

government is aware of this issue and is offering high salaries and research funding to attract talents. However, it mainly attracts Chinese nationals who have studied and worked overseas. Even though China's generous programmes have also targeted non-Chinese talents, the unclear and complex administrative procedures in visa and work permits application, as well as the lack of consistent policies, have created considerable constraints for them to work in China.⁴⁵



At a deeper level, the biggest challenge China faces is the lack of an environment conducive to making original breakthroughs. Chinese research institutions and companies tend to prioritise the development of AI applications and products according to market demands, and not on long-term investment in basic research. As noted by an industry observer, “so far there is literally no Chinese company investing heavily in fundamental research. No one is doing the step from 0 to 1”.⁴⁶ Yet only innovation from 0 to 1 is the basis for making a truly breakthrough innovation. This requires continuous efforts on theoretical research of fundamentals in AI.

On the other hand, the United States has been leading the world in fundamental research in AI.⁴⁷ For example, various deep learning frameworks including Convolutional Neural Networks, Recurrent Neural Networks and Transformers (upon which ChatGPT and Sora are based), as well as commonly used open-source software in AI development such as TensorFlow, have all been developed in the United States.

While China leads the world in publishing papers and patent applications, far surpassing that of the United States, it is still significantly behind in the most cutting-edge, fundamental and

⁴⁵ https://webapps.ilo.org/wcmsp5/groups/public/---asia/---ro-bangkok/---ilo-beijing/documents/publication/wcms_570674.pdf, accessed April 2024.

⁴⁶ <https://www.scmp.com/tech/tech-trends/article/3216241/chinas-chatgpt-quandary-result-lack-vision-and-fundamental-research-say-researchers>, accessed March 2024.

⁴⁷ <https://www.nus.edu.sg/newshub/news/2021/2021-12/2021-12-13/AI-lhzb-p12-13dec.pdf>, accessed March 2024.

EAI COMMENTARY

original research areas, becoming merely a follower of the United States. Therefore, the emergence of Sora once again prompts Chinese tech companies to ask themselves the soul-searching question of why breakthroughs occur elsewhere, rather than originating from China.⁴⁸

The Chinese government has also recognised this issue and begun investing heavily in basic science. China's research system is largely top-down, mobilising national resources and manpower under the leadership of the party and government to achieve goals. Such a mobilisation system may help China catch up rapidly if there is a clear goal in mind, such as developing AI models like Sora. However, whether such a system can nurture the next groundbreaking innovation like ChatGPT and Sora, which deeply transform human society, remains a subject of debate.

SHAN Wei is Senior Research Fellow of the East Asian Institute, National University of Singapore.

EAI commentaries serve to provide quick insights on current topics, based on ongoing research. The opinions expressed in the commentaries are those of the authors and do not necessarily reflect those of the East Asian Institute, or its Board Members.

EAI values your feedback and inputs ...

We would appreciate if you can spare a few minutes in giving us your feedback and comments on EAI Commentary No. **78** that you have just read.

Please visit <https://forms.office.com/r/bw5PZ0RdW3> to access a short survey form. Your inputs would be tremendously helpful to us in improving this series. Once again, thank you for your continuous support.

Best regards,
East Asian Institute,
National University of Singapore

⁴⁸ <https://www.scmp.com/tech/big-tech/article/3253034/openais-sora-pours-cold-water-chinas-ai-dreams-text-video-advancements-prompt-more-soul-searching>, accessed March 2024.