Trends in Food Science & Technology 148 (2024) 104488

Contents lists available at ScienceDirect

Trends in Food Science & Technology

journal homepage: www.elsevier.com/locate/tifs

Large language models in food science: Innovations, applications, and future

Peihua Ma[a], Shawn Tsai[b], Yiyang He[a], Xiaoxue Jia[a], Dongyang Zhen[c], Ning Yu[d], Qin Wang[a], Jaspreet K.C. Ahuja[a,e], Cheng-I Wei[a,*]

[a] Department of Nutrition and Food Science, College of Agriculture and Natural Resources, University of Maryland, College Park, MD, 20742, United States
[b] Beltsville Agricultural Research Center, Agriculture Research Service, United States Department of Agriculture, Beltsville, MD, 20705, United States
[c] Department of Civil and Environmental Engineering, A. James Clark School of Engineering, University of Maryland, College Park, MD, 20742, United States
[d] Salesforce Research, Palo Alto, CA, 94301, United States
[e] Beltsville Human Nutrition Research Center, Agriculture Research Service, United States Department of Agriculture, Beltsville, MD, 20705, United States

A R T I C L E I N F O

Keywords:
Natural language processing
Generative AI
Pre-trained model
Large language model

A B S T R A C T

Background: Large Language Models (LLMs) are increasingly significant in food science, transforming areas such as recipe development, nutritional analysis, food safety, and supply chain management. These models bring sophisticated decision-making, predictive analytics, and natural language processing capabilities to various aspects of food science.

Scope and approach: The review focuses on the application of LLMs in enhancing food science, with a strong emphasis on food safety, especially in contaminant detection and risk assessment. It addresses the roles of AI and LLMs in regulatory compliance and food quality control. Challenges like data biases, misinformation risks, and implementation hurdles, including data limitations and ethical concerns, are discussed. The necessity for interdisciplinary collaboration to overcome these challenges is also highlighted.

Key findings and conclusions: LLMs hold significant potential in automating processes and improving accuracy and efficiency in the global food system. Successful implementation requires continuous updates and ethical considerations. The paper provides insights for academics, industry professionals, and policymakers on the impact of LLMs in food science, emphasizing the importance of interdisciplinary efforts in this domain. Despite potential challenges, the integration of LLMs in food science promises transformative advancements.

1. Introduction

The field of food science and technology, integral to the advancement of global food systems, has historically been shaped by technological innovations. In this digital era, the emergence of Artificial Intelligence (AI) stands as a monumental development, ushering in a new epoch of data-driven decision-making and innovation. Among the various AI technologies, large language models (LLMs) such as the Generative Pre-trained Transformer (GPT) series have gained prominence due to their profound impact on data analysis and interpretation (Kasneci et al., 2023). These models, leveraging deep learning algorithms, have transformed the landscape of information processing, offering unparalleled insights into complex datasets (Carlini et al., 2021; Thirunavukarasu et al., 2023). In short, LLMs are a categories of models which can generate text by predicting the next token or word based on an input text. In the context of food science, GPT and similar LLMs are revolutionizing the field by enabling advanced natural language processing, predictive analytics, and enhanced decision-making capabilities. From optimizing recipes based on nutritional content to predicting food safety risks and improving supply chain efficiency, the applications of LLMs are diverse and continually expanding (Hong et al., 2020). This review paper seeks to explore the multifaceted role of GPT and other large language models in food science and technology, providing a comprehensive analysis of their influence, applications, and potential future developments (Liu et al., 2023).

This review articulates multiple objectives without categorizing itself as either a narrative or systematic review. Firstly, it aims to provide a detailed overview of how LLM, particularly the GPT series, are being integrated into various aspects of food science and technology. The paper examines a range of applications including, but not limited to,

* Corresponding author.
E-mail addresses: peihua@umd.edu (P. Ma), wei@umd.edu (C.-I. Wei).

https://doi.org/10.1016/j.tifs.2024.104488
Received 24 January 2024; Received in revised form 31 March 2024; Accepted 8 April 2024
Available online 16 April 2024

recipe development, nutritional analysis, food safety, and supply chain management. By collating and synthesizing existing literature, case studies, and empirical data, this review intends to offer a clear picture of the current state and future potential of LLMs in this field. The scope of this review is deliberately broad yet focused, covering the spectrum of LLM applications in food science while also addressing the ethical, technical, and practical challenges associated with their use. Furthermore, this review serves as a clarion call for more interdisciplinary collaboration between AI experts, food scientists, and industry stakeholders. Such collaborations are essential for harnessing the full potential of LLMs in addressing critical challenges in food science, from improving nutritional outcomes to ensuring food safety in an ever-changing global landscape. As such, the exploration of LLMs in food science is not just a study of a technological tool, but a reflection of the ongoing evolution of the field itself, adapting and advancing in response to new scientific and technological frontiers. Through this approach, the paper seeks to provide a valuable resource for academics, industry professionals, and policymakers interested in the intersection of AI and food science.

1.1. *Fundamentals of LLMs*

In the realm of computational linguistics, advanced models known as LLMs have emerged as pivotal tools. These models, exemplified by GPT-3 (Brown et al., 2020), Pathways Language Model (PaLM) (Chowdhery et al., 2023), and Large Language Model Meta AI (LLaMA, (Touvron, Lavril, et al., 2023), are distinguished from small language models by their colossal number of parameters, running into the hundreds of billions (Fig. 1). Trained on extensive text corpora, these models demonstrate remarkable proficiency in interpreting human language and executing complex text-based tasks.

The backbone of LLMs is the Transformer architecture, which utilizes deep networks with numerous attention mechanisms (Vaswani et al., 2017). While they share foundational aspects with their smaller counterparts, LLMs differentiate themselves through model complexity, data volume, and computational power on a significantly larger scale. This scaling has been proven to substantially augment the capabilities of these models (Radford et al., 2019). Post pre-training, LLMs possess a broad spectrum of capabilities applicable to a diverse array of tasks (A representative Post pre-training LLM chat system structure showed in Fig. 2a). To tailor these capabilities to specific requirements, a process known as instruction tuningis employed (Wei et al., 2021). This method involves fine-tuning the pre-trained LLMs with a curated set of instructions presented in natural language, thereby refining the models' abilities. The preparation for instruction tuning involves the collection and formulation of these instructional instances, which are then used to fine-tune the LLMs through supervised learning techniques, such as sequence-to-sequence loss. The outcome of this process is a marked improvement in the LLMs' capacity to generalize across novel tasks, demonstrating enhanced performance even in multilingual contexts (Muennighoff et al., 2022).

A notable aspect of LLMs Is the emerging novel capabilities that are not evident In smaller models (Wei, Tay, et al., 2022). This phenomenon becomes increasingly common when the models reach a certain scale. Among these emergent capabilities are: (1) The ability for contextual adaptation, as demonstrated by the 175B variant of GPT-3, enables these models to generate appropriate responses based on provided instructions and task demonstrations, bypassing the need for additional training. (2) Adaptability to instructions, where models are fine-tuned with diverse task datasets described in natural language, equipping them to handle new tasks described similarly (Ouyang et al., 2022; Sanh et al., 2021). (3) Sequential reasoning capability, which allows these

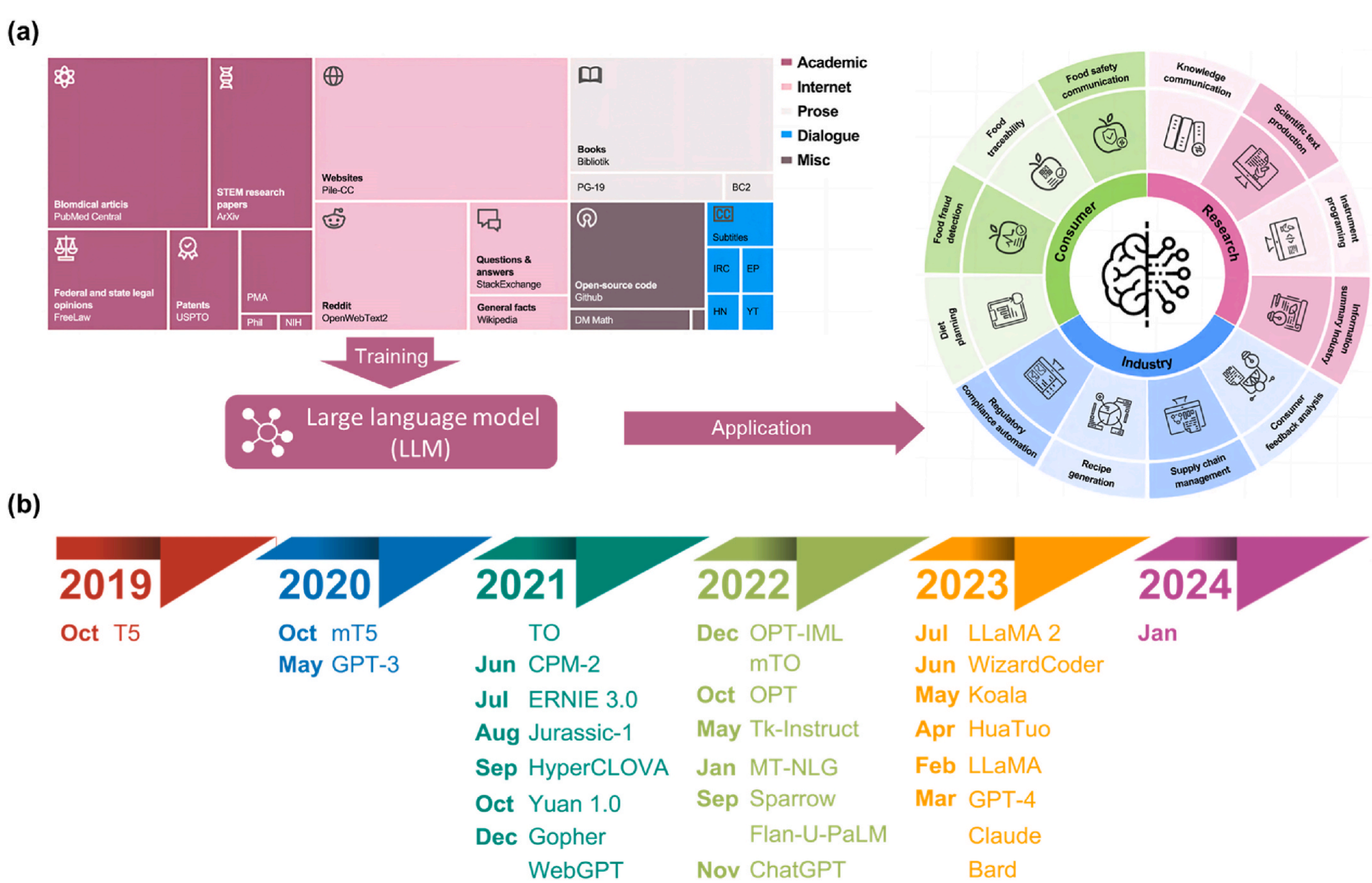

Fig. 1. (a) LLM training data sources categorized and displayed relative to their type. The application of LLMs in food related fields was summarized; (b) Milestones of LLM works.

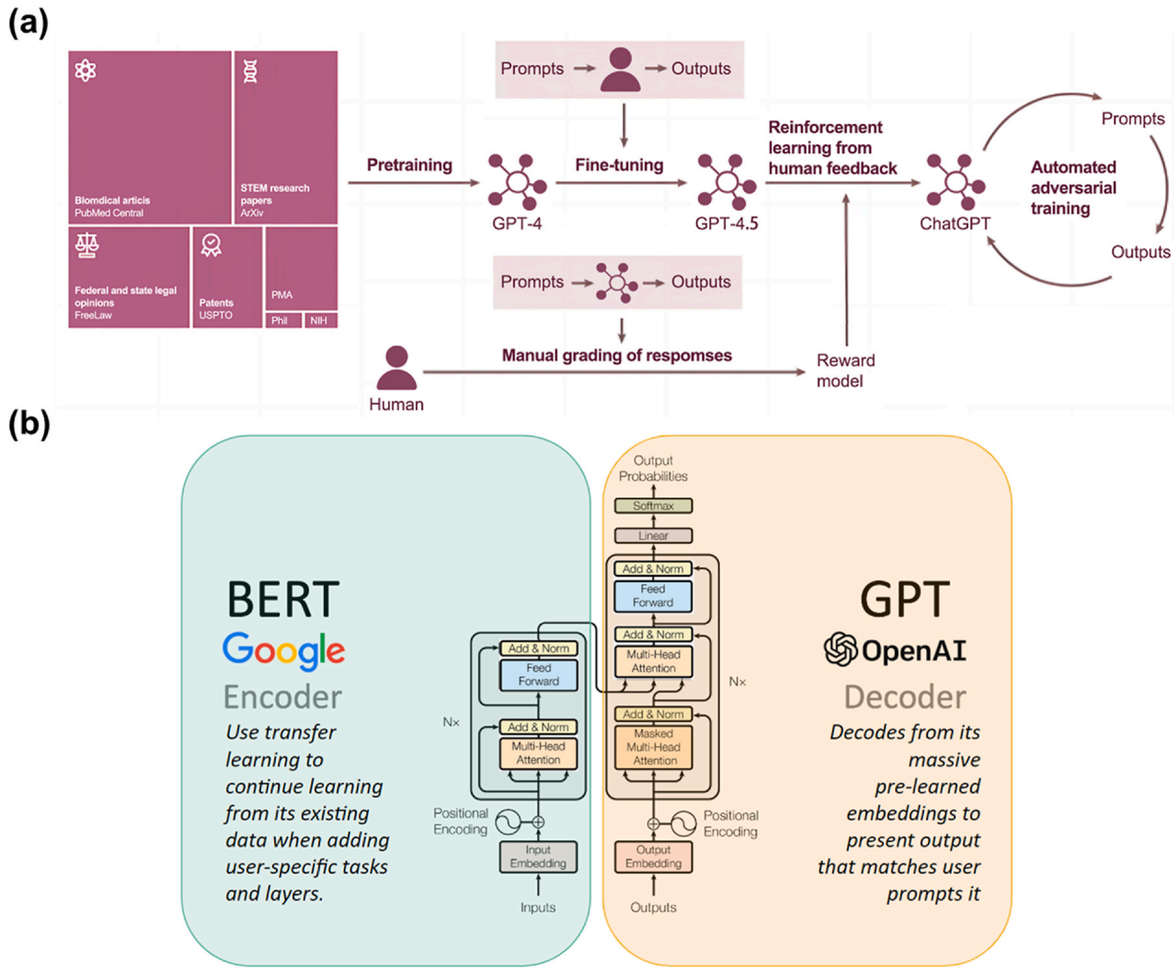

Fig. 2. (a) A typical application work flow of GPT where the prompts output was generated from fine-tuned GPT model. (b) The comparison between BERT and GPT. Although they both share the transformer architecture to learn context from textual-based datasets using attention mechanisms, BERT acts as an encoder while GPT acts as a decoder.

models to address tasks requiring multiple logical steps, such as complex mathematical problems (Wei, Wang, et al., 2022). This is achieved through a prompting strategy that includes intermediate steps in reasoning. The effectiveness of this strategy increases with the model's size and varies across different tasks.

The phenomenon of emergent abilities manifesting predominantly in LLMs rather than in smaller models presents a debating enigma (Gupta, Nair, Mishra, Ibrahim, & Bhardwaj, 2024). This observation underscores a broader issue: the absence of comprehensive, in-depth analyses identifying the critical factors that endow LLMs with their advanced capabilities. Understanding the conditions and mechanisms through which LLMs acquire these abilities is crucial for advancing our knowledge in the field. While there have been valuable discussions surrounding this topic, as highlighted in the works, there remains a pressing need for more systematic, theory-driven research efforts (Chen et al., 2024). Such investigations are essential to demystify the underlying principles that contribute to the remarkable performance of LLMs, paving the way for new insights and advancements in the domain of language modeling.

The evolution of LLMs to their current state of the art is attributed to several key developments: (1) The effectiveness of LLMs enhances with increased model size and data volume, coupled with greater computational resources (Hoffmann et al., 2022; Kaplan et al., 2020). Models like GPT-3 and PaLM have pushed these boundaries, showcasing the benefits of scaling. (2) Given their immense size, LLMs pose significant training challenges. This has led to the adoption of distributed training algorithms and various parallel computation strategies. Optimization techniques and frameworks like DeepSpeed (Rasley, Rajbhandari, Ruwase, & He, 2020) and Megatron-LM (Shoeybi et al., 2019) have been instrumental in this regard. (3) LLMs possess versatile abilities as general task solvers. Techniques like contextual prompting and fine-tuning with natural language descriptions are employed to harness these capabilities, especially in complex reasoning tasks (Fig. 2b). In summary, LLMs have marked a significant leap in language processing capabilities, driven by advancements in model scaling, training methodologies, and capability adaptation (Zhao et al., 2023).

1.2. Historical perspective

As we delve into the applications of LLMs within the food science field, it is essential to first understand the rich history and fast-pace progression of these models, which have revolutionized data processing and analysis in ways that are now beginning to impact food science research. The development of LLM has gone through four stages: statistical models, deep learning models, pre-training models and large-scale pre-training models (Naveed et al., 2023; Zhao et al., 2023). The origin of language modeling can be traced back to statistical models, particularly n-gram models. These laid the foundation for modern natural language processing (NLP), which is widely used in many tasks, such as speech recognition, handwriting recognition, spelling correction, machine translation, and search engines. The model estimates the occurrence probability of a given sentence in the corpus, extrapolating parameters by the Markov hypothesis and maximum likelihood estimation (Brown, DeSouza, Mercer, Della Pietra, & Lai, 1992). However, the reliance of n-gram models on statistics made it insufficient to catch deeper linguistic structures and long-range dependencies in languages.

With the introduction of neural networks, especially Recurrent Neural Networks (RNNs), language models are capable of processing sequential data, allowing for information retention across sequences, addressing a key limitation of n-gram models (Bengio, Ducharme, & Vincent, 2000; Mikolov, Karafiát, Burget, Cernocký, & Khudanpur, 2010). Another breakthrough came with the Transformer model in the paper "Attention Is All You Need" in 2017 (Vaswani et al., 2017). The Transformer model, leveraging parallel processing and multi-head attention mechanisms, offered a more efficient way to handle long-range dependencies, outperforming RNNs and LSTMs in scalability and effectiveness. This model set new benchmarks across various NLP tasks and became the backbone for subsequent advancements in language modeling.

LLMs refer to large-scale pre-trained language models. With the support of text corpus databases and self-supervised pre-train technology, LLMs offer powerful generic representation capabilities in solving complex problems (Naveed et al., 2023). In 2018, Google released Bidirectional Encoder Representations from Transformers (BERT), a model that utilizes the Autoencoding method for pre-training (Historical overview of LLM were list in Fig. 1b). BERT's bidirectional training, which considers both left and right context in all layers, enabled a deeper understanding of context in language, with the number of parameters exceeding 300 million for the first time (Devlin, Chang, Lee, & Toutanova, 2018). Google further developed this model into a more accurate version, RoBERTa, by using more training data, resources, and a dynamic mask adjustment strategy, achieving top performance in multiple NLP tasks (Liu et al., 2019).

Simultaneously, OpenAI introduced the GPT-1 model. GPT-1 was an autoregressive model, in contrast to BERT's autoencoding approach. The model assumes that the probability distribution of words at specific positions depends on all previous words, making its attention unidirectional (Radford, Narasimhan, Salimans, & Sutskever, 2018). Following this, OpenAI released GPT-2 in 2019. Adopting a similar architecture as GPT-1, the parameter size increased to 1.5 billion using the WebText dataset. GPT-2 conducts multi-task learning in the form of unsupervised language modeling, achieving good performance by expanding the model's capacity and data diversity (Radford et al., 2019). In 2020, OpenAI released the GPT-3 model, which is based on the extended model architecture of GPT-2 but uses over 175 billion parameters. GPT-3 introduces the concept of contextual learning to guide the model to understand tasks in the form of natural language text. Based on context learning, the pre-training goals and output of GPT-3 converge to the same language modeling paradigm: pre-training predicts text sequences based on context, and downstream tasks predict task solutions through context learning followed by formatting and output of text sequences (Brown et al., 2020). GPT-3 has shown excellent performance in multiple NLP tasks and in tasks that require reasoning or domain adaptation capabilities. ChatGPT, an AI dialogue system based on the GPT-3 model, generated widespread societal attention to AI technology after its release. The evolution continued with the launch of GPT-4, a large-scale, multimodal model which can accept image and text inputs and produce text outputs. GPT-4 leverages the Transformer model and has been fine-tuned with techniques like Reinforcement Learning from Human Feedback (RLHF), exhibiting human-level performance on professional and academic tasks (Fig. 2a). In 2023, Meta AI released the LLaMA model with versions ranging from 7 to 65 billion parameters, quickly becoming a popular open-source LLM (Touvron, Lavril, et al., 2023). LLaMA's lower computational cost made fine-tuning a mainstream approach for developing specialized models. The LLaMA-2 model, introduced in July 2023, used a larger and higher-quality training corpus and significantly improved dialogue generation (Touvron, Martin, et al., 2023; After July 2023, several new large language models (LLMs) were introduced, including Falcon 180B (Almazrouei et al., 2023) , Gemini (Team et al., 2023), DeepSeek-Coder (Guo et al., 2024), and DocLLM (Wang et al., 2023). Falcon 180B, developed by the Technology Innovation Institute (TII), is a highly capable open-source LLM with 180 billion parameters, trained on 3.5 trillion tokens. It is known for outperforming GPT-3.5 on the MMLU benchmark and is comparable to Google's PaLM 2-Large on various benchmarks. Gemini is a family of multimodal large language models developed by Google DeepMind, which serves as the successor to LaMDA and PaLM 2. It includes models like Gemini Ultra, Pro, and Nano, designed to handle complex tasks across different types of information such as text, code, audio, images, and video. Gemini has shown state-of-the-art performance on numerous benchmarks and is the first model to outperform human experts on the MMLU test. Representative LLM were summarized in Table 1.

2. Applications of LLMs in food science and technology

2.1. Predictive analytics for food safety

One of the most groundbreaking applications of LLMs lies in predictive analytics for contaminant detection and food safety risk assessment (Wang et al., 2022). Specifically, by analyzing vast datasets from past incidents, research studies, and safety reports, LLMs are capable of predicting potential contamination risks in food products (Deng, Cao, & Horn, 2021). This represents a significant shift from traditional microbiological and chemical testing methods to sophisticated computational analytics. Importantly, a notable example of this application involves a study investigating ChatGPT, a prominent LLM, in formulating diets for individuals with food allergies (Niszczota & Rybicka, 2023). The research critically assessed the model's ability to generate allergen-free diets, focusing on their safety, accuracy, and attractiveness. However, the study revealed limitations, particularly the model's potential to inadvertently include allergens and inaccuracies in calculating food quantities and energy values, thereby emphasizing the necessity for professional oversight. Consequently, such predictive capabilities are crucial, especially given the complexities of the globalized food supply chain (Geng, Zhao, Tao, & Han, 2017). Here, LLMs' advanced pattern recognition and anomaly detection capabilities can identify subtle risks that traditional methods might overlook. This preemptive risk identification by LLMs is invaluable, positioning them as proactive tools in preventing public health crises and as transformative elements in the realm of food safety analytics.

In applying LLMs to food safety, several innovative methodologies have emerged, notably integrating NLP and time-series forecasting (Makridis, Mavrepis, & Kyriazis, 2023). Makridis et al. (2023) utilized these technologies to demonstrate how LLMs can process extensive unstructured data and anticipate future food safety issues. Specifically, NLP techniques facilitate the extraction of meaningful insights from a

Table 1
Overview of LLM.

Model Name	Year	# of Parms	# of Tokens
BERT	2018	340M	137B
GPT-1	2018	1.3B	1.3B
GPT-2	2018	10B	10B
RoBERTa	2019	355M	2.2T
GPT-3	2020	175B	300B
BLOOM	2022	176B	366B
PaLM	2022	540B	780B
Falcon 180B	2023	46.7B	3.5T
Gemini	2023	3.25B	–
GPT-4	2023	1.76T	13T
LLaMA1	2023	65B	1.4T
LLaMA2	2023	70B	2T
DeepSeek-Coder	2024	33B	2T
DocLLM	2024	7B	2T

wide range of textual data associated with food safety, including regulatory documents, recall notices, and research publications. In addition, these results are effectively translated into predictions about future food safety incidents in combination with time-series forecasting. The combination of NLP and time-series forecasting enables proactive rather than reactive management. However, there are still limitations in analyzing spatial and temporal data, which is critical in the food safety area, using NLP and time-series prediction alone. Thus, researchers have turned to deep learning models like Convolutional Neural Networks (CNNs) and RNNs to unravel the complex patterns often seen in food recall and contamination events, which have shown a remarkable ability to provide highly accurate forecasts (Armghan et al., 2023). Consequently, this may has the potential that allows stakeholders to incorporate LLMs into the food supply chain effectively, enabling them to take timely and targeted actions.

Indeed, the realm of predictive models in food safety extends well beyond just LLMs. Notably, Artificial Neural Networks (ANNs) have demonstrated their potential in areas such as modeling microbial growth and improving predictions related to food safety. However, ANNs face specific challenges, including a tendency for slow learning during their training phase and a susceptibility to becoming trapped in local optima (Chen, Li, Dou, Wen, & Dong, 2023). These limitations can significantly impact their overall efficiency and the accuracy of their predictions. Another challenge with ANNs lies in the complexity of interpreting how these models arrive at their conclusions, raising concerns about their transparency and applicability in crucial food safety decision-making scenarios (Cheng, Huang, Ruess, & Yasuoka, 2018). Meanwhile, the adoption of LLMs in food safety also presents its own set of challenges, including the need for high-quality data, substantial computational resources, and enhancements in model interpretability (Tamkin, Brundage, Clark, & Ganguli, 2021). Despite these hurdles, the distinctive capability of LLMs to process complex, unstructured data clearly sets them apart, underlining their potential to bring transformative changes in food safety predictive analytics. In conclusion, the growing integration of LLMs in the field of food safety, as exemplified in the provided case study, is primarily driven by their unparalleled ability to interpret and analyze a diverse array of data sources. This proficiency is essential for preemptively identifying risks and thus plays a critical role in enhancing public health within an ever-evolving global food landscape.

2.2. Regulatory compliance automation

Regulatory compliance automation involves leveraging technology to ensure adherence to legal standards and regulations (Oguejiofor et al., 2023). This is essential for maintaining public trust and ensuring product quality. The dynamic and global nature of regulations in the food industry presents a significant challenge. Keeping abreast of varied laws and standards in different areas requires considerable effort (Makofske, 2019). Conventional methods often involve manual documentation, along with human-led inspections and audits (Henson & Heasman, 1998). These traditional processes also require individuals to manually analyze regulations, a method that is slow and prone to errors. Consequently, producers, who must continuously adapt to the dynamic nature of food safety regulations, have increased risks of non-compliance. Due to the complex and ever-changing nature of regulations, innovative approaches are required to ensure compliance and minimize the risk of penalties (Garcia Martinez, Verbruggen, & Fearne, 2013). LLMs, which can automate the analysis and interpretation of extensive regulatory documents, thus have emerged as crucial tools in this context (Arbel & Becher, 2023). These models are adept at understanding legal terminologies and provide concise summaries and actionable insights for compliance. This advancement not only streamlined the regulatory process but also improved accuracy in adherence to regulations. These models excel in parsing, analyzing, and offering actionable insights from extensive regulatory texts. For example, Wu, Liu, Wu, Chen, and Tong (2021) demonstrated the efficacy of deep learning-based NLP models in the pharmaceutical domain by successfully classifying drug-induced liver injury risk in FDA labeling documents (Wu et al., 2021). This study illustrates the capability of LLMs to precisely navigate and interpret complex regulatory language, a skill highly transferable to the food safety domain. Furthermore, Ershov (2023) showcased how language models can be employed to automate the construction of knowledge graphs for compliance (Ershov, 2023). These graphs translate intricate legal texts into structured, comprehensible formats, thereby simplifying the compliance process. Such an approach is directly applicable to managing the European Union's Food Safety and Quality regulations, offering a streamlined method for food producers to remain compliant with evolving standards. Additionally, the research by Winter, van der Aa, Rinderle-Ma, and Weidlich (2020) revealed how NLP can facilitate the comparison and alignment of regulatory documents with process model repositories (Winter et al., 2020). This helped ensure that business processes were in line with the latest regulatory requirements, crucial for maintaining compliance in the dynamic field of food safety.

However, the use of LLMs in regulatory compliance is not without challenges. Ensuring the accuracy and relevance of the information provided by LLMs is a key concern (Thapa & Adhikari, 2023). While proficient in processing large text volumes, they may not always capture the nuances of regulatory changes, influenced by cultural, regional, or political factors. There is a risk of bias in training data, which could bias insights (Wu, Duan, & Ni, 2023). Despite these challenges, LLMs in regulatory compliance processes offer a promising path forward. Automating tedious tasks can significantly reduce time and resource investment. Nonetheless, it is essential to complement these technological solutions with human expertise for nuanced understanding and judgment. In conclusion, LLMs present a transformative opportunity in regulatory compliance automation. Their successful application depends on a balanced approach that combines technological efficiency with human oversight (Ghimire, Kim, & Acharya, 2023). As the field evolves, LLMs are likely to become an indispensable part of compliance strategies in the food industry and beyond.

2.3. Enhancing inspection and quality control

Incorporating LLMs into inspection processes has the potential to significantly enhance the accuracy and comprehensiveness of quality control measures in the food industry. While specific case studies are limited, the theoretical applications of these advanced models are promising. LLMs, with their sophisticated capabilities in NLP, can be utilized to analyze extensive datasets from quality control processes (Viellieber & Aßenmacher, 2020). This could include interpreting and correlating findings from various sources such as textual reports, sensor data, and even social media trends to identify subtle patterns and differences that might indicate quality issues or emerging risks (Zhang et al., 2022). Consider a conceptual application of LLMs in food quality control. It could be used to continuously monitor online consumer feedback (Mohagheghi & Aagedal, 2007). The model can analyze quality control reports as they are generated. In real-time, it would identify any trends indicating a potential decline in product quality. It could also spot early signs of contamination (Velichety, Ram, & Bockstedt, 2019). Using pattern recognition and difference detection, this approach allows for timely responses to emerging quality issues and enables manufacturers and distributors to address potential issues swiftly (Lee, Sung, & Jeon, 2019). As a result, food safety can be maintained to high standards and waste can be reduced. Such applications, though currently theoretical, highlight the transformative potential of LLMs in revolutionizing food quality inspection and control. The efficiency of LLMs in tasks like semantic classification and multi-step logical reasoning supports their potential in such applications (Creswell, Shanahan, & Higgins, 2022). Additionally, their ability to translate natural language goals into structured planning languages and generate comprehensive long-form answers further demonstrates their versatility and utility in complex reasoning and analysis tasks, which are crucial for quality control in the food industry (Chen, 2022; Xie et al., 2023).

A specialized LLM for Food Quality Control, designed for the food testing domain, demonstrates the immense potential of LLMs in revolutionizing inspection and quality control within the food industry (Chowdhery et al., 2023; Qi, Yu, Tu, Tan, & Huang, 2023). This conceptual model is notable for its incremental pre-training on a diverse array of data sources, such as images, scanned documents, and structured data from private databases (Fig. 3a). This approach enables it to adeptly assimilate and interpret the complex and varied information typical in food testing scenarios (Xu et al., 2020). A key feature of this LLM is the integration of a knowledge graph, significantly enhancing the accuracy of its outputs by mitigating instances of machine hallucination, a common challenge in LLM applications (Ross, Martinez, Houde, Muller, & Weisz, 2023). This precision is particularly crucial in food testing, where accurate data interpretation can have significant implications for food safety and quality (Lin et al., 2022). Despite its advanced capabilities, this specialized LLM faces challenges, notably the need for specific, high-quality training data (Workshop et al., 2022). This requirement mandates continuous updates and refinements to align with the evolving standards and practices in food testing, presenting scalability and adaptability challenges, especially when expanding its application to different types of food products or testing protocols (Antkiewicz, Czarnecki, & Stephan, 2009). Nonetheless, by effectively processing and analyzing diverse datasets, the LLM aids in identifying patterns and differences that could indicate quality issues, enabling earlier interventions to maintain high food safety standards. This early detection capability is invaluable in reducing waste and ensuring consumer trust. Its development and deployment illustrate both the vast capabilities and the practical considerations involved in utilizing LLMs for complex, real-world applications in food science.

The use of LLMs in the areas of inspection and quality control is poised to bring about significant improvements (Sadilek et al., 2018). Such models, capable of processing and analyzing extensive data from various sources, stand to enhance the precision and thoroughness of food quality and safety evaluations (Huang, Kangas, & Rasco, 2007). LLMs have the unique ability to unearth subtle patterns and irregularities in data, which might otherwise go unnoticed by human inspectors. This capability is crucial for the early identification of possible quality issues, thereby enabling prompt and effective responses. Such early detection is essential in averting food spoilage and safeguarding consumer health (Torres et al., 2022). Additionally, LLM for Food Quality Control incorporates knowledge graphs to significantly boost the model's capacity to generate accurate and dependable outputs (Fig. 3b). This is particularly vital in the exacting realm of food testing, where

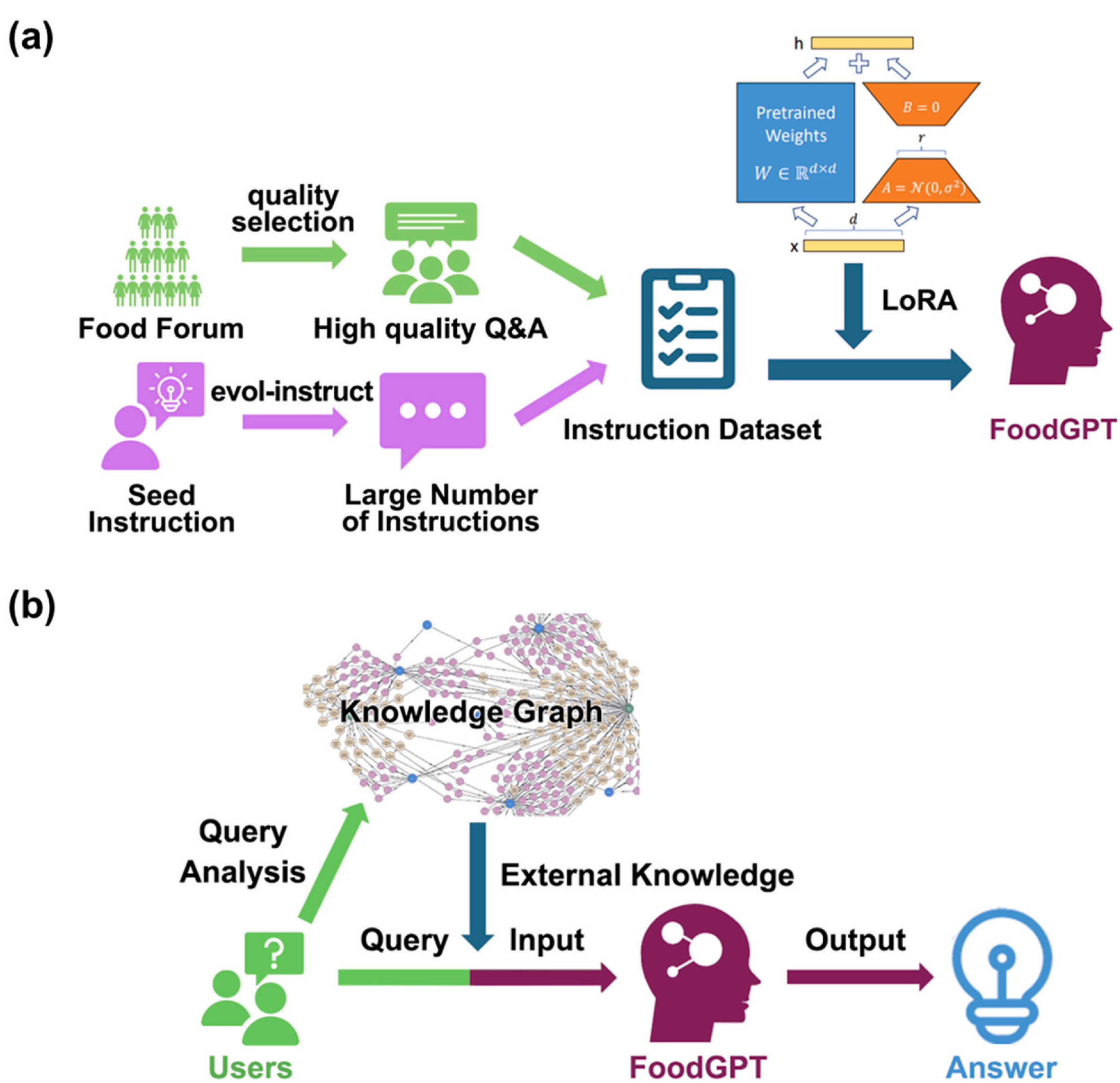

Fig. 3. (a) The entire pipeline of instruction fine-tuning of a FoodGPT; (b) Retrieval method for FoodGPT from external knowledge graph.

precision is non-negotiable (Bro et al., 2002). The fusion of comprehensive, accurate data interpretation with these models ensures their insights and recommendations are reliably informed and relevant. In summary, the deployment of LLMs in the processes of inspection and quality control within food science marks a notable progression towards smart and autonomous food safety inspection. These models are instrumental in elevating the quality of safety and quality assessments, thereby contributing to the elevation of food safety standards. As these technologies evolve, their integration into the food industry's strategies for ensuring product safety and quality is expected to become increasingly central.

2.4. Consumer feedback analysis for quality improvement

The ability of LLMs to analyze unstructured data such as customer reviews and feedback has been invaluable for quality improvement. One work that integrates insights into consumer feedback analysis in food science reveals an important advancement (Ouyang et al., 2022). It illustrates that fine-tuning language models like InstructGPT with human feedback can lead to a significant improvement in understanding and responding to consumer needs in the food industry. By aligning these models more closely with user intentions, the potential for accurate and beneficial insights in product development and quality improvement is greatly enhanced. This approach highlights the importance of tailoring language models to comprehend and interpret nuanced consumer sentiments and preferences, thereby making them more effective and relevant in the evolving landscape of food science.

Another work focusing on the challenges and evolution of integrating human feedback into LLMs is particularly relevant for understanding consumer preferences and values in food products (Kirk, Bean, Vidgen, Röttger, & Hale, 2023). It emphasizes the importance of developing LLMs that are sensitive and responsive to diverse human tastes and cultural nuances, which is crucial for food companies aiming to tailor their products to meet the varying and subjective preferences of their consumer base. This approach underscores the need for more nuanced, inclusive, and culturally aware models in analyzing consumer feedback, ensuring that product development and innovation in the food industry are both consumer-centric and culturally sensitive.

Moreover, one paper introduces an innovative approach for making AI predictions in language models, like BERT, more understandable and transparent (Binder, Heinrich, Hopf, & Schiller, 2022). It utilizes linguistic rules to globally reconstruct the model's predictions, thereby enhancing their explainability. This is particularly relevant for online consumer reviews in e-commerce, where understanding the basis of AI-driven analysis is crucial. This approach balances fidelity and comprehensibility in AI, contributing to more transparent and accountable AI applications in areas like sentiment analysis and text analytics. This methodology could be highly beneficial in the food science domain, where understanding consumer reviews through a transparent and explainable AI system can lead to more insightful and reliable product improvements and customer engagement strategies. Furthermore, another study emphasizes the importance of high-quality feedback in enhancing the performance of language models (Cui et al., 2023). It introduces the concept of UltraFeedback, which uses detailed feedback, both numerical and textual, to refine language models. This methodology is crucial in the context of food science, particularly for analyzing consumer feedback. Implementing such sophisticated feedback mechanisms can significantly improve the accuracy and relevance of insights derived from consumer reviews, leading to more effective product development and innovation in the food industry. The use of UltraFeedback exemplifies how detailed, quality feedback can elevate the performance of language models in practical applications.

The use of LLMs in analyzing consumer feedback enables businesses to gain deeper insights into consumer preferences, tailor their products more effectively, and respond proactively to market trends (Alamsyah & Girawan, 2023; Kamal & Himel, 2023). As LLM technology continues to evolve, its application in consumer feedback analysis is set to become increasingly sophisticated, offering unprecedented opportunities for quality improvement and customer satisfaction in the food industry.

2.5. Supply chain management and traceability

LLMs are also revolutionizing supply chain management and traceability in the food industry. The incorporation of the Internet of Things (IoT) into supply chain management significantly augments the capabilities of LLMs in ensuring food safety and traceability. The application of the Bayes Classifiers Algorithm, coupled with IoT, provides a more robust system for tracking and tracing food products throughout the supply chain (Balamurugan, Ayyasamy, & Joseph, 2019). This method ensures an efficient and secure transfer of food products from manufacturers to consumers, reinforcing the food supply chain's resilience against disruptions and contamination risks. This integration exemplifies the progressive move towards more intelligent, data-driven approaches in managing the complexities of food supply chain logistics.

Recent research has promoted an innovative approach, exemplified by OptiGuide, which revolutionizes supply chain management by bridging the gap between complex supply chain data and human comprehension (Li, Mellou, Zhang, Pathuri, & Menache, 2023). By translating intricate supply chain optimization solutions into understandable insights, OptiGuide enhances decision-making efficiency and accuracy in the industry. The integration of LLMs with optimization technology in this manner promises to significantly advance supply chain management, offering a more dynamic, responsive, and transparent approach to handling the intricate networks that define modern food supply chains. Other research adds depth to the understanding of how NLP can transform supply chain management (Schöpper & Kersten, 2021; Singla, Anandayuvaraj, Kalu, Schorlemmer, & Davis, 2023). It highlights the potential of NLP in extracting complex supply chain information from diverse data sources, providing a more comprehensive view of the entire supply network. LLMs, as advanced implementations of NLP, are capable of processing and interpreting extensive supply chain data, offering nuanced insights and improved transparency. This approach is particularly beneficial for enhancing supply chain visibility, a crucial factor in managing today's complex and globalized supply chains. However, the study also underscores the need for further research and development in this area to overcome current challenges and fully realize the potential of NLP in supply chain optimization. Within the context of Industry 4.0, another paper offers an insightful examination of NLP techniques in presenting the latest developments in supply chain management (Zhou, Awasthi, & Stal-Le Cardinal, 2021). This study emphasizes the growing importance of integrating NLP with supply chain operations, particularly in the analysis of complex multi-tier structures. By leveraging NLP, supply chains can efficiently process and analyze complex datasets, facilitating a deeper understanding of multi-tier supply networks. The application of NLP in this context allows for the extraction of valuable insights from vast amounts of unstructured data, thereby enhancing decision-making processes and optimizing supply chain operations. This integration not only improves operational efficiency but also ensures greater adaptability and resilience in the face of rapidly changing market demands and technological advancements.

2.6. Food fraud detection

In recent years, the detection and prevention of food fraud has become a critical concern in the food industry. With their advanced analytical capabilities, LLMs can process vast amounts of data from various sources, such as supply chain records, online marketplaces, and regulatory databases, to identify patterns and inconsistencies indicative of food fraud (Ooi et al., 2023).

One study focuses on the MedISys-FF tool, which utilizes global media data to identify trends and developments in food fraud activities

(Marvin et al., 2022). The tool has shown efficacy in collecting and analyzing data from a vast range of countries and in multiple languages. This approach enhances the capability to detect food fraud early and may assist all stakeholders in the food system to ensure safe, healthy, and authentic food. The integration of such media monitoring tools with LLMs can significantly enhance the ability to detect and respond to food fraud incidents quickly and efficiently. This method represents a promising direction in the use of advanced technology for ensuring food safety and integrity in the global food supply chain.

The application of LLMs in food fraud detection involves analyzing labels, certification documents, and testing results. By processing this information, LLMs can detect discrepancies that may suggest adulteration, mislabeling, or use of unauthorized substances in food products. This capability is particularly important in a globalized food market where the complexity of supply chains can often obscure fraudulent activities. One related study explores a unique combination of deep learning and NLP in identifying and standardizing fake food images (Mezgec, Eftimov, Bucher, & Seljak, 2019). This innovative approach aids in automated dietary assessment, utilizing a deep learning model for precise image recognition, followed by NLP methods for food matching and standardization. This technique is significant for its potential application in food fraud detection, where accurately identifying and categorizing food items is crucial. By extending this methodology to real food products, it can significantly enhance the detection and prevention of food fraud in the food industry, contributing to safer and more reliable food supply chains while offering a more efficient way to ensure food authenticity and safety.

Recent research demonstrates how a federated Bayesian network (BN) model, integrating data from various sources without compromising data privacy, can effectively predict food fraud types (Gavai et al., 2023). This approach addresses challenges in data sharing due to privacy concerns, making it a powerful tool for collaborative food fraud detection and prevention in the food supply chain. By leveraging a federated BN model, LLMs can now analyze data from multiple sources while maintaining data privacy and security. This method not only enhances the detection accuracy of food fraud but also fosters collaboration and trust among different actors in the supply chain. This innovative approach aligns with the growing need for sophisticated, privacy-preserving techniques in combating food fraud in the food industry.

However, the effectiveness of LLMs in food fraud detection is contingent upon the quality and comprehensiveness of the data they analyze. Incomplete or biased data can lead to inaccurate predictions or overlooked fraudulent activities (Okada, Mertens, Liu, Lam, & Ong, 2023). Therefore, continuous updating and training of LLMs with diverse and extensive datasets are crucial for maintaining their accuracy and reliability. While challenges remain, the potential of LLMs to enhance food safety and protect consumers is immense. As these models continue to evolve, they will become even more integral to combat food fraud.

Table 2
Summary of LLMs application in recipe generation and modification.

Name	Aims	Results	References
Skip-gram Models	To assist in choosing alternative ingredients or recipes and generate new recipes with authentic flavors	Generated new recipes with a higher number of recognizable cooking steps and ingredients, especially using LSTM models	Pan et al. (2020)
ReProg and ReComp Approaches	To generate creative culinary recipes using language models and genetic algorithms	Achieved better recipe generation results, particularly in the utilization of secondary ingredients like oils and seasonings	(Antô, Bezerra, Góes, & Ferreira, 2020)
RecipeGPT	To generate and evaluate recipes using a generative pre-trained transformer (GPT-2)	Effectively generated both ingredient texts and cooking instruction texts, showing comparable performance to human authors	(H. Lee et al., 2020)
Word Embedding-Based Method	To adapt ingredient recipes to user preferences in an unsupervised manner	Successfully adapted recipes to fit dietary needs and preferences, achieving high satisfaction rates in user surveys	(Morales-Garzón et al., 2021)
RecipeGM	To generate recipes using hierarchical convolutional neural networks with self-attention mechanisms	Outperformed LSTM and RecipeGPT in some cases, discussing the need for suitable metrics for NLG in recipe generation	Reusch et al. (2021)
Ratatouille	To generate new realistic cooking recipes using deep learning models, particularly LSTMs and GPT-2	The GPT-2 model outperformed LSTM models in maintaining contextual information and producing results closer to actual reference texts	(Goel et al., 2022)
FOODGPT	To reduce food waste by generating personalized recipes based on user preferences and available ingredients	Enabled users to substitute ingredients effectively, capturing the nuances in relationships and interactions of multiple ingredients	Qi et al. (2023)
Cook-Gen	To extend food computation models for a comprehensive understanding of recipes and cooking actions	Reliably generated cooking actions from recipes, outperforming other large language models and baselines	Venkataramanan et al. (2023)
Large Language Models as Sous Chefs	To adapt and simplify complex culinary recipes into more accessible formats	Showed a preference for revised recipes over the original ones, enhancing recipe accessibility and understanding	(Hwang, Li, Hou, & Roth, 2023)
FIRE	To generate comprehensive recipes including titles, ingredients, and cooking instructions from food images	Successfully created detailed recipes by integrating vision and language models	Chhikara et al. (2024)
Nutritional analysis and diet planning			
HANA	Predicting death status based on food categories during COVID-19	Predicted death status with high accuracy, aligned with WHO's nutrition recommendations	Shams et al. (2021)
AgriBERT	Matching food descriptions with nutritional data in agriculture	Outperformed other language models in semantic matching of food and nutrition	Rezayi et al. (2022)
ChatGPT	Generating meal plans for non-communicable diseases patients	Improved meal planning accuracy with additional information, though further improvement needed	(Papastratis, Stergioulas, Konstantinidis, Daras, & Dimitropoulos, 2024)
ChatGPT and GPT 4.0	AI nutritionist program for patients with type 2 diabetes mellitus	Showed high accuracy in recognizing ingredients and aligning food recommendations with expert advice	Sun et al. (2023)
KG4NH	Comprehensive knowledge graph for dietary nutrition and human health	High precision, recall, and F1 score in nutrition and disease relation extraction	(Fu et al., 2023)
NMT-RL	Teacher-forced REINFORCE for nutritionally enhanced diets	Outperformed both traditional and modern approaches in diet planning for children	Lee, Kim, Lim, et al. (2021)

2.7. Recipe generation and modification

The development of LLMs for recipe generation and modification is an intricate process, beginning with pre-training on diverse culinary datasets like Recipe1M, Recipe NLG, Recipe Ingredients, Food.com, Beer Recipes, Epicurious, and Recipe Box, which contain a large number of recipes, ingredients, and consumer feedback (Katserelis & Skianis, 2022). These models adeptly interpret a variety of inputs, from dish names to images, utilizing lexical analysis and contextual understanding to generate comprehensive recipes (summarized in Table 2). For textual inputs, the model recognizes cuisine type, flavor profiles, and cooking methods to refine recipes, while for image-based inputs, advanced computer vision techniques like CNN or Vision Transformer (ViT) analyze dish images to extract visual features (Chhikara, Chaurasia, Jiang, Masur, & Ilievski, 2024). These features are then translated into embeddings and matched against a database of recipe embeddings to identify close counterparts. Finally, the model synthesizes this information to produce detailed recipes, including ingredient lists, cooking instructions, and even alternative ingredients suggestions (Pan, Xu, & Li, 2020). This workflow epitomizes the nuanced capabilities of generative AI in transforming diverse inputs into structured, innovative culinary outputs, significantly enhancing the domain of digital gastronomy.

One of the notable methodologies is Food Image to Recipe Generation (FIRE), which employs a multimodal approach combining vision and language models to generate comprehensive recipes from food images. FIRE utilizes the BLIP model for title generation, ViT for ingredient extraction, and T5 Model for cooking instruction creation. This methodology represents a significant leap forward in generating accurate and detailed recipes from just one image, illustrating the power of combining different AI technologies.

The adaptation of recipes to specific dietary needs or preferences is another area where LLMs excel. Models utilizing Word2Vec for ingredient embeddings facilitate ingredient substitutions, offering alternatives that cater to individual dietary restrictions and preferences (Cunningham, Farnum, Kang, & Saxena, 2023). Cook-Gen (Venkataramanan et al., 2023), a generative AI approach, extends the understanding of recipes beyond ingredients and nutrition to include cooking actions, enhancing the adaptability and interaction with recipes. These adaptations are critical in personalizing recipes, making them more accessible and aligned with user needs.

In the realm of recipe adaptation, a word embedding-based method for unsupervised adaptation of cooking recipes offers an innovative approach (Morales-Garzón, Gómez-Romero, & Martin-Bautista, 2021). It proposes an unsupervised method to adapt ingredient recipes to user preferences by creating and applying a domain-specific word embedding model. This model aims for quality adapted recipes without prior knowledge extraction that addresses user needs and preferences, including nutritional preferences and diet restrictions like vegetarian and vegan diets. The model's effectiveness is highlighted by the positive feedback from users who evaluated adapted recipes.

The Skip-gram Model is another method to help choose alternative ingredients or recipes and generate new recipes with authentic flavors of certain cuisine styles (Pan et al., 2020). The model uses word embeddings to understand how ingredients are combined and substituted and build n-gram and neural network models for generating new recipes. This model's effectiveness is evident in the higher number of recognizable cooking steps and ingredients in the recipes generated, particularly using LSTM models.

Further developments in this field are showcased by tools such as RecipeGM and RecipeGPT (Reusch, Weber, Thiele, & Lehner, 2021). RecipeGM uses hierarchical CNNs with self-attention mechanisms for recipe generation, demonstrating an ability to generate diverse and nuanced recipes. RecipeGPT, utilizing generative pre-trained transformers and fine-tuned on extensive cooking recipe datasets, showcases effectiveness in generating coherent recipes that align with user intent and contextual accuracy (H. Lee et al., 2020).

An additional tool contributing to the culinary field is RecipeScape, an interactive system designed for analyzing cooking instructions at scale (Chang et al., 2018). It assists culinary professionals and students in categorizing cooking methods and understanding ingredient usage patterns. This system exemplifies the potential of LLMs to provide in-depth analysis and insights, contributing to educational and professional pursuits in the culinary arts.

2.8. Nutritional analysis and diet planning

The integration of AI in the field of nutrition and dietetics aims to enhance the accuracy, personalization, and efficiency of dietary assessments and interventions. In particular, LLMs, a subset of AI, show promise in understanding complex dietary data, providing personalized nutrition advice, and aiding in diet planning (Sak & Suchodolska, 2021).

In the scope of nutritional analysis, a variety of studies have elucidated the expansive capabilities of LLMs and deep learning techniques. The HANA model, developed for assessing the impact of dietary habits on COVID-19 outcomes, exemplifies the use of regression systems in predicting health outcomes based on food consumption patterns. By identifying dietary components correlated with morbidity and mortality rates, this model underscores the significant role nutrition plays in disease outcomes, guided by data from a comprehensive COVID-19 Healthy Diet Dataset encompassing global food and health statistics (Shams et al., 2021). Similarly, AgriBERT harnesses a domain-specific corpus of food-related and agricultural literature, exhibiting the power of fine-tuned LLMs in the semantic matching of food descriptions to nutritional data, significantly outperforming general models and baselines in this complex task (Rezayi et al., 2022).

Moreover, advancements in AI for nutrient estimation from food images and ingredient statements have been notable. Kaur et al., systematically reviewed deep learning's role in classifying food images and identifying nutrient content (Kaur, Kumar, & Gupta, 2023). Similarly, deep learning approaches have been employed to predict food categories and nutrients, utilizing vast databases like the USDA's Branded Food Products Database (BFPD). These models, which include multilayer perceptron (MLP), RNN, CNN, and graph neural network (GNN), have shown up to 99% accuracy in food classification and remarkable precision in nutrient estimation, emphasizing the potential of AI in automating and refining the accuracy of tasks within the food and nutrient domain (Ma et al., 2021). Furthermore, the ChinaMartFood-109 study explores deep learning for precision dietary records based on images, utilizing CNNs for food recognition and nutrient estimation. The CNNs achieved notable accuracy and demonstrate the efficacy of AI in food analysis, particularly from image data (Ma, Lau, Yu, Li, & Sheng, 2022).

Diet planning is complex, necessitating an understanding of nutritional needs, cultural eating patterns, and individual preferences. LLMs have been employed to assist in generating diet plans that align with specific health conditions and preferences. Studies involving AI dietitians for T2DM management reveal how LLMs can synthesize vast amounts of dietary and health data to provide tailored advice, catering to the unique nutritional needs and health conditions of individuals. This level of customization extends to addressing specific health conditions, like Type 2 Diabetes Mellitus, where diet plays a crucial role in management and prognosis (Sun et al., 2023). The MIcrosoft News Dataset (MIND) dataset represents another significant leap in diet planning, providing a rich, structured compilation of menus, ingredients, nutrients, and diet patterns to facilitate machine learning applications in nutrition (Lee et al., 2021). This comprehensive dataset allows for more nuanced and accurate machine learning models that can better understand and predict dietary needs and preferences.

Additionally, interactive systems such as EZNutriPal are at the forefront of engaging users in their dietary planning process. By employing a human-in-the-loop approach, these systems continuously learn and adapt to individual feedback, enhancing the personalization of diet monitoring and advice. The iterative learning process inherent in these models ensures that dietary recommendations are not static but evolve in response to changing health metrics, preferences, and lifestyle factors, thus offering a dynamic, user-centric approach to diet planning and nutritional management (Hezarjaribi et al., 2019). Such advancements underscore the potential of LLMs to not only understand and analyze complex nutritional data, but to also actively engage and empower individuals in their dietary and health journeys.

Furthermore, the integration of AI and LLMs in nutritional science is vividly demonstrated through studies focused on assessing nutrient intake for hospitalized patients and predicting nutritional biomarkers. A study utilizes RGB-Depth (RGB-D) image pairs for a novel approach to accurately estimate nutrient intake. This system's integration of a multi-task contextual network, a few-shot learning classifier, and a 3D surface construction algorithm markedly improves the accuracy and efficiency of nutritional assessments in hospital settings. It represents a significant leap forward in managing disease-related malnutrition and enhancing the precision of dietary interventions for patient care. On the other hand, deep learning was used in another study to predict serum levels of pyridoxal 5′-phosphate, a critical marker of vitamin B6 status. The model used in the study demonstrates a significant enhancement over traditional multivariable linear regression models. This breakthrough illustrates the potential of AI in capturing complex dietary and biochemical data to refine nutritional assessment and research. These examples underscore the transformative impact of AI and LLMs in nutritional science, offering a glimpse into the future of personalized, accurate, and efficient dietary management and nutritional research. Together, they highlight the broader implications of these technologies in advancing healthcare and nutrition understanding, paving the way for future research and personalized nutritional interventions.

3. Challenges and future perspectives

3.1. Challenges

Data limitations and bias LLMs predominantly rely on datasets that over-represent popular or extensively documented cuisines, leading to underrepresentation of less-known culinary practices. This skewed data distribution affects the model's proficiency in providing diverse and culturally nuanced food-related information. Furthermore, evolving food safety standards and emerging research findings are often not immediately incorporated into LLM training datasets, creating a gap between current knowledge and the model's information base.

Bias in LLMs is multifaceted, encompassing cultural, linguistic, and research-focused aspects while linguistic bias is rooted in the predominance of English-language data, marginalizing non-English food terminologies and practices. Additionally, food safety often deals with unique or rare cases, such as specific food allergies or rare pathogens. LLMs may not handle these well due to limited exposure in their training data. The inherent bias in scientific research towards well-funded or widely studied topics results in an uneven representation of food safety issues in the training data. This bias is further compounded by the publication bias towards studies with significant findings, leading to an incomplete portrayal of the spectrum of scientific research in the models' dataset. Addressing these biases necessitates a deliberate effort to diversify training data, regularly update models with new research, and implement measures to mitigate existing biases. However, quantifying the extent of these biases in LLMs, specifically in the context of food safety and science, is an area that requires more comprehensive research.

Risk of misinformation The risk of misinformation is a paramount concern, primarily due to the models' reliance on their training datasets. Datasets containing inaccuracies, biases, or outdated information predispose LLMs to perpetuate such misinformation in their outputs. The absence of real-time data validation further exacerbates this issue, as LLMs cannot incorporate recent scientific advancements or changes in food safety regulations post-training. Additionally, LLMs tend to generalize complex subjects, a tendency that is particularly problematic in the nuanced field of food safety, where oversimplifications can lead to erroneous advice. The implications of misinformation in this context are grave, especially considering the potential health risks. Incorrect guidance on food handling, preparation, or storage can lead to serious consequences like foodborne illnesses. Misinformation in nutrition advice can adversely impact individuals with specific dietary requirements or health conditions. Furthermore, such misinformation can influence public perception and policy decisions, possibly leading to non-evidence-based practices.

To mitigate these risks, several strategies are crucial. Ensuring regular updates and validation of the training data in alignment with the latest scientific findings is essential. Involving domain experts in reviewing and validating the outputs of LLMs, particularly for critical advice, can enhance accuracy. Educating users about the inherent limitations of LLMs and the importance of consulting additional, reliable sources for crucial decisions is also vital. Finally, employing advanced training methodologies, such as fine-tuning models with high-quality, peer-reviewed scientific literature, can significantly reduce the likelihood of misinformation and improve the overall reliability of LLMs in the domain of food safety and science.

Language and terminology challenges LLMs face significant challenges in accurately representing the linguistic diversity and technical specificity inherent in the field of food safety and science. The predominance of English-language data in LLM training leads to a linguistic bias, underrepresenting non-English terms and cultural specificities related to food. This issue is compounded by the wide array of terminologies unique to various cuisines and food practices, as well as regional variations within the same language. Additionally, the technical jargon of food science presents a barrier, requiring LLMs to not only comprehend but also accurately translate complex scientific terminology into more accessible languages, a task often hindered by the evolving nature of scientific language and concepts.

The accurate interpretation and translation of food-related terms are further challenged by the nuances of context and the cultural significance attached to certain foods and practices. Misinterpretation can lead to inaccuracies and cultural insensitivity, particularly when handling traditional foods and customs. To mitigate these issues, a comprehensive approach is needed. This includes diversifying linguistic and cultural representation in training datasets, continuously updating models with new terms and concepts, and integrating expert input for more accurate interpretations. Additionally, incorporating training focused on cross-cultural sensitivity can enhance the LLMs' capacity to respect and accurately convey the cultural context of food-related language. In essence, addressing language and terminology challenges in LLMs requires a concerted effort towards linguistic inclusivity, contextual accuracy, and cultural sensitivity, particularly vital in the specialized and culturally rich domain of food safety and science. And although some hallucination still exist in LLMs, understanding key words is an advantage of LLMs compared with traditional methods, especially if LLMs connect to external databases by tools like LangChain or after fine-tuning with high quality data in this field.

Ethical and Legal Considerations Ethical and legal considerations are of paramount importance due to the potential impact on public health, cultural integrity, and legal compliance. Ethically, the imperative lies in ensuring the accuracy and safety of the information provided, given the serious health consequences of misinformation. Moreover, LLMs must navigate cultural sensitivities with care, respecting diverse food practices and avoiding biases or stereotypes that could lead to cultural misrepresentation or offense. Privacy concerns also emerge, particularly when personalizing dietary advice, necessitating strict adherence to ethical standards of data privacy and user consent.

Legally, LLMs in the food domain must align with complex and varied food safety regulations, a task fraught with challenges given the international spectrum of these laws. Inaccurate or non-compliant advice could lead to health risks and subsequent legal liabilities.

Additionally, there is the legal aspect of intellectual property, especially pertinent when dealing with proprietary recipes or cooking methods with cultural significance. The potential liability for damages resulting from misinformation further complicates the legal landscape, raising questions about responsibility allocation among LLM creators, hosting platforms, and users.

To effectively address these ethical and legal challenges, a multifaceted approach is required. This includes rigorous testing and validation of LLM outputs, comprehensive cultural and ethical training to minimize biases, regular updates to ensure legal compliance, and transparent communication with users about the capabilities and limitations of LLMs. Such measures are crucial in maintaining the reliability, cultural appropriateness, and legal soundness of LLM applications in the sensitive and impactful field of food safety and science.

Difficulty of integration The integration of LLMs with other technologies in the food safety and science domain faces several limitations that need careful consideration. One significant challenge is data privacy and security concerns. Integrating LLMs with technologies like IoT devices and mobile applications involves handling sensitive consumer data, which raises questions about data protection and user privacy. Ensuring that this integration complies with stringent data privacy laws and maintains the confidentiality of user information is crucial.

Another key limitation is the issue of interoperability and standardization. For effective integration, it is essential that different technologies, such as IoT, image recognition, and blockchain, operate seamlessly with LLMs. This requires not only technical compatibility but also standardization of data formats and protocols. Achieving this interoperability is a complex task, necessitating collaboration among a wide range of stakeholders, including technology developers, food industry experts, and regulatory bodies.

Building user trust and acceptance presents an additional challenge. Users, whether they are consumers or professionals in the food industry, must have confidence in the accuracy and reliability of the integrated system. This is particularly important when the system combines LLM outputs with data from other technologies, such as real-time monitoring from smart kitchen appliances or image analysis for food quality assessments. Gaining user trust requires demonstrating the system's efficacy and reliability through transparent communication and robust validation processes.

A notable area of potential in this integration is AI-powered analytics for food science research. However, the integration of LLMs with AI-driven analytics platforms is not without limitations. The primary challenge lies in ensuring the quality and representativeness of the data fed into these platforms. AI analytics rely heavily on the availability of large, diverse datasets, and any bias or gaps in these datasets can significantly impact the insights generated. Moreover, translating complex AI analytics into actionable and understandable advice for food safety and science requires sophisticated processing and interpretation capabilities in LLMs. Overcoming these limitations necessitates rigorous data curation, continuous model updating, and the involvement of domain experts to ensure that the integrated system effectively supports advanced research in food science.

3.2. Perspectives

Emerging research in LLMs is poised to significantly enhance their application in food science and research. A key focus is on improving contextual understanding and memory. By advancing LLMs' ability to maintain and comprehend extended contextual information, they can more effectively process complex inquiries and literature pertinent to food science. This development promises to enhance the depth and coherence of LLM responses to intricate queries, crucial for both academic research and practical applications in food science.

Another notable direction is the enhancement of LLMs with multimodal capabilities, integrating text, images, and voice data. Such multimodal LLMs can offer comprehensive analyses by simultaneously interpreting scientific imagery and textual information. This integration is particularly beneficial for food science, where visual data, such as molecular structures or nutritional charts, play a crucial role. Moreover, the focus on domain adaptation tailors LLMs to specific fields, enhancing their proficiency in understanding and conveying technical food science terminology and concepts. This specialized adaptation significantly elevates the accuracy and relevance of LLMs in professional and research contexts within food science.

The commitment to mitigating biases and ensuring ethical AI use also stands to positively impact LLM applications in food science. By striving for fairness and representativeness, LLMs can provide more balanced and culturally sensitive insights, essential for the diverse and global nature of food practices and research. Additionally, the pursuit of advanced personalization techniques in LLMs will enable more customized dietary and nutritional advice, aligning closely with individual or group-specific needs in food research. Collectively, these emerging research directions promise not only to augment LLMs' functional capabilities, but also to profoundly enrich their utility and reliability in the dynamic and multifaceted realm of food science and research.

4. Conclusion

In summary, this review has elucidated the significant role of LLMs in revolutionizing various facets of food science. By leveraging their advanced capabilities in NLP, predictive analytics, and decision-making, LLMs offer unprecedented opportunities for innovation in recipe development, nutritional analysis, food safety, and supply chain efficiency. However, their application is not without challenges. Issues such as data biases, the potential for misinformation, and the necessity for ethical considerations and continuous model updates are critical concerns that need to be addressed. The integration of LLMs in regulatory and quality control processes presents a promising avenue for enhancing accuracy and efficiency, yet it demands careful implementation and oversight. The future of LLMs in food science relies heavily on interdisciplinary collaboration and ongoing research to realize their full potential while mitigating associated risks. As we continue to explore and expand the boundaries of these technologies, their contribution to advancing global food systems remains a vital area of exploration for researchers, industry professionals, and policymakers alike.

Data availability

No data was used for the research described in the article.

References

Alamsyah, A., & Girawan, N. D. (2023). Improving clothing product quality and reducing waste based on consumer review using RoBERTa and BERTopic Language Model. *Big Data and Cognitive Computing, 7*(4), 168.

Antkiewicz, M., Czarnecki, K., & Stephan, M. (2009). Engineering of framework-specific modeling languages. *IEEE Transactions on Software Engineering, 35*(6), 795–824.

Antô, W., Bezerra, J. R., Góes, L. F. W., & Ferreira, F. M. F. (2020). Creative culinary recipe generation based on statistical language models. *IEEE Access, 8*, 146263–146283.

Arbel, Y. A., & Becher, S. (2023). How smart are smart readers? llms and the future of the no-reading problem. *LLMs and the Future of the No-Reading Problem*.

Armghan, A., Logeshwaran, J., Sutharshan, S. M., Aliqab, K., Alsharari, M., & Patel, S. K. (2023). Design of biosensor for synchronized identification of diabetes using deep learning. *Results in Engineering, 20*, Article 101382.

Balamurugan, S., Ayyasamy, A., & Joseph, K. (2019). An efficient bayes classifiers algorithm for traceability of food supply chain management using internet of things. *International Journal of Engineering and Advanced Technology, 9*(1), 2995–3005.

Bengio, Y., Ducharme, R., & Vincent, P. (2000). A neural probabilistic language model. *Advances in Neural Information Processing Systems, 13*.

Binder, M., Heinrich, B., Hopf, M., & Schiller, A. (2022). Global reconstruction of language models with linguistic rules–Explainable AI for online consumer reviews. *Electronic Markets, 32*(4), 2123–2138.

Bro, R., van den Berg, F., Thybo, A., Andersen, C. M., Jørgensen, B. M., & Andersen, H. (2002). Multivariate data analysis as a tool in advanced quality monitoring in the food production chain. *Trends in Food Science & Technology, 13*(6–7), 235–244.

Brown, P., DeSouza, P., Mercer, R., Della Pietra, V., & Lai, J. (1992). Class-based n-gram models of natural language. *Computational Linguistics*, 1950.
Brown, T., Mann, B., Ryder, N., Subbiah, M., Kaplan, J. D., Dhariwal, P., … Askell, A. (2020). Language models are few-shot learners. *Advances in Neural Information Processing Systems, 33*, 1877–1901.
Carlini, N., Tramer, F., Wallace, E., Jagielski, M., Herbert-Voss, A., Lee, K., … Erlingsson, U. (2021). Extracting training data from large language models. In *30th USENIX security symposium (USENIX security 21)*.
Chang, M., Guillain, L. V., Jung, H., Hare, V. M., Kim, J., & Agrawala, M. (2018). Recipescape: An interactive tool for analyzing cooking instructions at scale. In *Proceedings of the 2018 CHI conference on human factors in computing systems*.
Chen, W. (2022). Large language models are few (1)-shot table reasoners. *arXiv preprint arXiv:2210.06710*.
Chen, Y., Li, H., Dou, H., Wen, H., & Dong, Y. (2023). Prediction and visual analysis of food safety risk based on TabNet-GRA. *Foods, 12*(16), 3113.
Chen, J., Liu, Y., Li, D., An, X., Feng, Z., Zhao, Y., et al. (2024). Plug-and-Play grounding of reasoning in multimodal large language models. *arXiv preprint arXiv:2403.19322*.
Cheng, C.-H., Huang, C.-H., Ruess, H., & Yasuoka, H. (2018). Towards dependability metrics for neural networks. In *2018 16th ACM/IEEE international conference on formal methods and models for system design (MEMOCODE)*.
Chhikara, P., Chaurasia, D., Jiang, Y., Masur, O., & Ilievski, F. (2024). Fire: Food image to REcipe generation. In *Proceedings of the IEEE/CVF winter conference on applications of computer vision*.
Chowdhery, A., Narang, S., Devlin, J., Bosma, M., Mishra, G., Roberts, A., … Gehrmann, S. (2023). Palm: Scaling language modeling with pathways. *Journal of Machine Learning Research, 24*(240), 1–113.
Creswell, A., Shanahan, M., & Higgins, I. (2022). Selection-inference: Exploiting large language models for interpretable logical reasoning. *arXiv preprint arXiv:2205.09712*.
Cui, G., Yuan, L., Ding, N., Yao, G., Zhu, W., Ni, Y., … Sun, M. (2023). *Ultrafeedback: Boosting language models with high-quality feedback. arXiv preprint arXiv:2310.01377*.
Cunningham, R., Farnum, S., Kang, S., & Saxena, R. (2023). *Foodgpt: AMachine learning approach to ingredient substitution and recipe recommendation*.
Deng, X., Cao, S., & Horn, A. L. (2021). Emerging applications of machine learning in food safety. *Annual Review of Food Science and Technology, 12*, 513–538.
Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2018). Bert: Pre-training of deep bidirectional transformers for language understanding. *arXiv preprint arXiv:1810.04805*.
Ershov, V. (2023). A case study for compliance as code with graphs and language models: Public release of the regulatory knowledge graph. *arXiv preprint arXiv:2302.01842*.
Fu, C., Pan, X., Wu, J., Cai, J., Huang, Z., van Harmelen, F., … He, T. (2023). KG4NH: A Comprehensive Knowledge Graph for Question Answering in Dietary Nutrition and Human Health. *IEEE Journal of Biomedical and Health Informatics*.
Garcia Martinez, M., Verbruggen, P., & Fearne, A. (2013). Risk-based approaches to food safety regulation: What role for co-regulation? *Journal of Risk Research, 16*(9), 1101–1121.
Gavai, A., Bouzembrak, Y., Mu, W., Martin, F., Kaliyaperumal, R., van Soest, J., … Marvin, H. J. (2023). Applying federated learning to combat food fraud in food supply chains. *Npj Science of Food, 7*(1), 46.
Geng, Z., Zhao, S., Tao, G., & Han, Y. (2017). Early warning modeling and analysis based on analytic hierarchy process integrated extreme learning machine (AHP-ELM): Application to food safety. *Food Control, 78*, 33–42.
Ghimire, P., Kim, K., & Acharya, M. (2023). Generative AI in the construction industry: Opportunities & challenges. *arXiv preprint arXiv:2310.04427*.
Goel, M., Chakraborty, P., Ponnaganti, V., Khan, M., Tatipamala, S., Saini, A., & Bagler, G. (2022). *May). Ratatouille: A tool for Novel Recipe Generation* (pp. 107–110). IEEE.
Gupta, R., Nair, K., Mishra, M., Ibrahim, B., & Bhardwaj, S. (2024). Adoption and impacts of generative artificial intelligence: Theoretical underpinnings and research agenda. *International Journal of Information Management Data Insights, 4*(1), Article 100232.
Lee, H., Shu, K., Achananuparp, P., Prasetyo, P. K., Liu, Y., Lim, E.-P., et al. (2020). RecipeGPT: Generative pre-training based cooking recipe generation and evaluation system. In *Companion proceedings of the web conference 2020*.
Henson, S., & Heasman, M. (1998). Food safety regulation and the firm: Understanding the compliance process. *Food Policy, 23*(1), 9–23.
Hezarjaribi, N., Mazrouee, S., Hemati, S., Chaytor, N. S., Perrigue, M., & Ghasemzadeh, H. (2019). Human-in-the-loop learning for personalized diet monitoring from unstructured mobile data. *ACM Transactions on Interactive Intelligent Systems (TiiS), 9*(4), 1–24.
Hoffmann, J., Borgeaud, S., Mensch, A., Buchatskaya, E., Cai, T., Rutherford, E., … Clark, A. (2022). Training compute-optimal large language models. *arXiv preprint arXiv:2203.15556*.
Hong, T., Kim, D., Ji, M., Hwang, W., Nam, D., & Park, S. (2020). *Bros: A pre-trained language Model for understanding texts in document*.
Huang, Y., Kangas, L. J., & Rasco, B. A. (2007). Applications of artificial neural networks (ANNs) in food science. *Critical Reviews in Food Science and Nutrition, 47*(2), 113–126.
Kamal, M., & Himel, A. S. (2023). Redefining modern marketing: An analysis of AI and NLP's influence on consumer engagement, strategy, and beyond. *Eigenpub Review of Science and Technology, 7*(1), 203–223.
Kaplan, J., McCandlish, S., Henighan, T., Brown, T. B., Chess, B., Child, R., … Amodei, D. (2020). *Scaling laws for neural language models. arXiv preprint arXiv:2001.08361*.
Kasneci, E., Seßler, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F., … Hüllermeier, E. (2023). ChatGPT for good? On opportunities and challenges of large language models for education. *Learning and Individual Differences, 103*, Article 102274.
Katserelis, K., & Skianis, K. (2022). Towards fine-dining recipe generation with generative pre-trained transformers. *arXiv preprint arXiv:2209.12774*.
Kaur, R., Kumar, R., & Gupta, M. (2023). Deep neural network for food image classification and nutrient identification: A systematic review. *Reviews in Endocrine and Metabolic Disorders, 24*(4), 633–653.
Kirk, H. R., Bean, A. M., Vidgen, B., Röttger, P., & Hale, S. A. (2023). The past, present and better future of feedback learning in large language models for subjective human preferences and values. *arXiv preprint arXiv:2310.07629*.
Lee, C., Kim, S., Jeong, S., Lim, C., Kim, J., Kim, Y., et al. (2021). MIND dataset for diet planning and dietary healthcare with machine learning: Dataset creation using combinatorial optimization and controllable generation with domain experts. In *Thirty-fifth conference on neural information processing systems datasets and benchmarks track (round 2)*.
Lee, S. W., Sung, H. J., & Jeon, H. M. (2019). Determinants of continuous intention on food delivery apps: Extending UTAUT2 with information quality. *Sustainability, 11*(11), 3141.
Li, B., Mellou, K., Zhang, B., Pathuri, J., & Menache, I. (2023). *Large language models for supply chain optimization. arXiv preprint arXiv:2307.03875*.
Lin, H., Wang, F., Duan, Y., Kang, W., Chen, Q., & Xue, Z. (2022). Early detection of wheat Aspergillus infection based on nanocomposite colorimetric sensor and multivariable models. *Sensors and Actuators B: Chemical, 351*, Article 130910.
Liu, Y., Han, T., Ma, S., Zhang, J., Yang, Y., Tian, J., … Liu, Z. (2023). *Summary of chatgpt-related research and perspective towards the future of large language models.* Meta-Radiology, Article 100017.
Liu, Y., Ott, M., Goyal, N., Du, J., Joshi, M., Chen, D., … Stoyanov, V. (2019). *Roberta: A robustly optimized bert pretraining approach. arXiv preprint arXiv:1907.11692*.
Ma, P., Lau, C. P., Yu, N., Li, A., Liu, P., Wang, Q., et al. (2021). Image-based nutrient estimation for Chinese dishes using deep learning. *Food Research International*, Article 110437.
Ma, P., Lau, C. P., Yu, N., Li, A., & Sheng, J. (2022). Application of deep learning for image-based Chinese market food nutrients estimation. *Food Chemistry, 373*, Article 130994.
Makofske, M. P. (2019). Inspection regimes and regulatory compliance: How important is the element of surprise? *Economics Letters, 177*, 30–34.
Makridis, G., Mavrepis, P., & Kyriazis, D. (2023). A deep learning approach using natural language processing and time-series forecasting towards enhanced food safety. *Machine Learning, 112*(4), 1287–1313.
Marvin, H. J., Bouzembrak, Y., Van der Fels-Klerx, H., Kempenaar, C., Veerkamp, R., Chauhan, A., … Vrolijk, H. (2022). Digitalisation and Artificial Intelligence for sustainable food systems. *Trends in Food Science & Technology, 120*, 344–348.
Mezgec, S., Eftimov, T., Bucher, T., & Seljak, B. K. (2019). Mixed deep learning and natural language processing method for fake-food image recognition and standardization to help automated dietary assessment. *Public Health Nutrition, 22*(7), 1193–1202.
Mikolov, T., Karafiát, M., Burget, L., Cernocký, J., & Khudanpur, S. (2010). *Recurrent neural network based language model. Interspeech*.
Mohagheghi, P., & Aagedal, J. (2007). *Evaluating quality in model-driven engineering. International workshop on modeling in software engineering (MISE'07: ICSE workshop 2007)*.
Morales-Garzón, A., Gómez-Romero, J., & Martin-Bautista, M. J. (2021). A word embedding-based method for unsupervised adaptation of cooking recipes. *IEEE Access, 9*, 27389–27404.
Naveed, H., Khan, A. U., Qiu, S., Saqib, M., Anwar, S., Usman, M., … Mian, A. (2023). *A comprehensive overview of large language models. arXiv preprint arXiv:2307.06435*.
Niszczota, P., & Rybicka, I. (2023). The credibility of dietary advice formulated by ChatGPT: Robo-diets for people with food allergies. *Nutrition, 112*, Article 112076.
Oguejiofor, B. B., Omotosho, A., Abioye, K. M., Alabi, A. M., Oguntoyinbo, F. N., Daraojimba, A. I., et al. (2023). A review on data-driven regulatory compliance in Nigeria. *International Journal of Applied Research in Social Sciences, 5*(8), 231–243.
Okada, Y., Mertens, M., Liu, N., Lam, S. S. W., & Ong, M. E. H. (2023). AI and machine learning in resuscitation: Ongoing research, new concepts, and key challenges. *Resuscitation, 15*, Article 100435.
Ooi, K.-B., Tan, G. W.-H., Al-Emran, M., Al-Sharafi, M. A., Capatina, A., Chakraborty, A., … Lee, V.-H. (2023). The potential of generative artificial intelligence across disciplines: Perspectives and future directions. *Journal of Computer Information Systems*, 1–32.
Ouyang, L., Wu, J., Jiang, X., Almeida, D., Wainwright, C., Mishkin, P., … Ray, A. (2022). Training language models to follow instructions with human feedback. *Advances in Neural Information Processing Systems, 35*, 27730–27744.
Pan, Y., Xu, Q., & Li, Y. (2020). Food recipe alternation and generation with natural language processing techniques. In *2020 IEEE 36th international conference on data engineering workshops (ICDEW)*.
Papastratis, I., Stergioulas, A., Konstantinidis, D., Daras, P., & Dimitropoulos, K. (2024). Can ChatGPT provide appropriate meal plans for NCD patients? *Nutrition, 121*, 112291.
Qi, Z., Yu, Y., Tu, M., Tan, J., & Huang, Y. (2023). *Foodgpt: A large language model in food testing domain with incremental pre-training and knowledge graph prompt*. arXiv preprint arXiv:2308.10173.
Radford, A., Narasimhan, K., Salimans, T., & Sutskever, I. (2018). *Improving language understanding by generative pre-training*.
Radford, A., Wu, J., Child, R., Luan, D., Amodei, D., & Sutskever, I. (2019). Language models are unsupervised multitask learners. *OpenAI blog, 1*(8), 9.
Rasley, J., Rajbhandari, S., Ruwase, O., & He, Y. (2020). Deepspeed: System optimizations enable training deep learning models with over 100 billion parameters. In *Proceedings of the 26th ACM SIGKDD international conference on knowledge discovery & data mining*.

Reusch, A., Weber, A., Thiele, M., & Lehner, W. (2021). In *RecipeGM: A hierarchical recipe generation model. 2021 IEEE 37th international conference on data engineering workshops (ICDEW)*.
Rezayi, S., Liu, Z., Wu, Z., Dhakal, C., Ge, B., Zhen, C., … Li, S. (2022). Agribert: Knowledge-infused agricultural language models for matching food and nutrition. In *Proceedings of the thirty-first international joint conference on artificial intelligence*.
Ross, S. I., Martinez, F., Houde, S., Muller, M., & Weisz, J. D. (2023). The programmer's assistant: Conversational interaction with a large language model for software development. In *Proceedings of the 28th international conference on intelligent user interfaces*.
Sadilek, A., Caty, S., DiPrete, L., Mansour, R., Schenk, J. T., Bergtholdt, M., … Gabrilovich, E. (2018). Machine-learned epidemiology: Real-time detection of foodborne illness at scale. *NPJ Digital Medicine, 1*(1), 36.
Sak, J., & Suchodolska, M. (2021). Artificial intelligence in nutrients science research: A review. *Nutrients, 13*(2), 322.
Sanh, V., Webson, A., Raffel, C., Bach, S. H., Sutawika, L., Alyafeai, Z., … Raja, A. (2021). Multitask prompted training enables zero-shot task generalization. *arXiv preprint arXiv:2110.08207*.
Schöpper, H., & Kersten, W. (2021). Using natural language processing for supply chain mapping: A systematic review of current approaches. In *5th international conference on computational linguistics and intelligent systems (COLINS 2021)*.
Shams, M. Y., Elzeki, O. M., Abouelmagd, L. M., Hassanien, A. E., Abd Elfattah, M., & Salem, H. (2021). Hana: A healthy artificial nutrition analysis model during COVID-19 pandemic. *Computers in Biology and Medicine, 135*, Article 104606.
Shoeybi, M., Patwary, M., Puri, R., LeGresley, P., Casper, J., & Catanzaro, B. (2019). Megatron-lm: Training multi-billion parameter language models using model parallelism. *arXiv preprint arXiv:1909.08053*.
Singla, T., Anandayuvaraj, D., Kalu, K. G., Schorlemmer, T. R., & Davis, J. C. (2023). An empirical study on using large language models to analyze software supply chain security failures. In *Proceedings of the 2023 workshop on software supply chain offensive research and ecosystem defenses*.
Sun, H., Zhang, K., Lan, W., Gu, Q., Jiang, G., Yang, X., … Han, D. (2023). An AI dietitian for type 2 Diabetes Mellitus management based on large language and image recognition models: Preclinical concept validation study. *Journal of Medical Internet Research, 25*, Article e51300.
Tamkin, A., Brundage, M., Clark, J., & Ganguli, D. (2021). Understanding the capabilities, limitations, and societal impact of large language models. *arXiv preprint arXiv:2102.02503*.
Thapa, S., & Adhikari, S. (2023). ChatGPT, bard, and large language models for biomedical research: Opportunities and pitfalls. *Annals of Biomedical Engineering, 51* (12), 2647–2651.
Thirunavukarasu, A. J., Ting, D. S. J., Elangovan, K., Gutierrez, L., Tan, T. F., & Ting, D. S. W. (2023). Large language models in medicine. *Nature Medicine, 29*(8), 1930–1940.
Torres, J. P., Caro, A., Ávila, M. D. M., Pérez-Palacios, T., Antequera, T., & Rodríguez, P. G. (2022). A computer-aided inspection system to predict quality characteristics in food technology. *IEEE Access, 10*, 71496–71507.
Touvron, H., Lavril, T., Izacard, G., Martinet, X., Lachaux, M.-A., Lacroix, T., … Azhar, F. (2023). *Llama: Open and efficient foundation language models. arXiv preprint arXiv: 2302.13971*.
Touvron, H., Martin, L., Stone, K., Albert, P., Almahairi, A., Babaei, Y., … Bhosale, S. (2023). Llama 2: Open foundation and fine-tuned chat models. *arXiv preprint arXiv: 2307.09288*.
Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., … Polosukhin, I. (2017). Attention is all you need. *Advances in Neural Information Processing Systems, 30*.
Velichety, S., Ram, S., & Bockstedt, J. (2019). Quality assessment of peer-produced content in knowledge repositories using development and coordination activities. *Journal of Management Information Systems, 36*(2), 478–512.
Venkataramanan, R., Roy, K., Raj, K., Prasad, R., Zi, Y., Narayanan, V., et al. (2023). Cook-Gen: Robust generative modeling of cooking actions from recipes. *arXiv preprint arXiv:2306.01805*.
Viellieber, V. D., & Aßenmacher, M. (2020). *Pre-trained language models as knowledge bases for Automotive Complaint Analysis. arXiv preprint arXiv:2012.02558*.
Wang, X., Bouzembrak, Y., Lansink, A. O., & van der Fels-Klerx, H. (2022). Application of machine learning to the monitoring and prediction of food safety: A review. *Comprehensive Reviews in Food Science and Food Safety, 21*(1), 416–434.
Wei, J., Tay, Y., Bommasani, R., Raffel, C., Zoph, B., Borgeaud, S., … Metzler, D. (2022). *Emergent abilities of large language models. arXiv preprint arXiv:2206.07682*.
Wei, J., Wang, X., Schuurmans, D., Bosma, M., Xia, F., Chi, E., … Zhou, D. (2022). Chain-of-thought prompting elicits reasoning in large language models. *Advances in Neural Information Processing Systems, 35*, 24824–24837.
Winter, K., van der Aa, H., Rinderle-Ma, S., & Weidlich, M. (2020). In *Assessing the compliance of business process models with regulatory documents. Conceptual modeling: 39th international conference, ER 2020, Vienna, Austria, November 3–6, 2020, proceedings 39*.
Workshop, B., Scao, T. L., Fan, A., Akiki, C., Pavlick, E., Ilić, S., … Yvon, F. (2022). *Bloom: A 176b-Parameter Open-Access Multilingual Language Model. arXiv preprint arXiv:2211.05100*.
Wu, X., Duan, R., & Ni, J. (2023). Unveiling security, privacy, and ethical concerns of chatgpt. *Journal of Information and Intelligence*. In press.
Wu, Y., Liu, Z., Wu, L., Chen, M., & Tong, W. (2021). BERT-based natural language processing of drug labeling documents: A case study for classifying drug-induced liver injury risk. *Frontiers in Artificial Intelligence, 4*, Article 729834.
Xie, Y., Yu, C., Zhu, T., Bai, J., Gong, Z., & Soh, H. (2023). *Translating natural language to planning goals with large-language models. arXiv preprint arXiv:2302.05128*.
Xu, P., Patwary, M., Shoeybi, M., Puri, R., Fung, P., Anandkumar, A., et al. (2020). MEGATRON-CNTRL: Controllable story generation with external knowledge using large-scale language models. *arXiv preprint arXiv:2010.00840*.
Zhang, S., Roller, S., Goyal, N., Artetxe, M., Chen, M., Chen, S., … Lin, X. V. (2022). Opt: Open pre-trained transformer language models. *arXiv preprint arXiv:2205.01068*.
Zhao, W. X., Zhou, K., Li, J., Tang, T., Wang, X., Hou, Y., … Dong, Z. (2023). A survey of large language models. *arXiv preprint arXiv:2303.18223*.
Zhou, R., Awasthi, A., & Stal-Le Cardinal, J. (2021). The main trends for multi-tier supply chain in Industry 4.0 based on Natural Language Processing. *Computers in Industry, 125*, Article 103369.
Wei, J., Bosma, M., Zhao, V. Y., Guu, K., Yu, A. W., Lester, B., & Le, Q. V. (2021). *Finetuned language models are zero-shot learners*.
Muennighoff, N., Wang, T., Sutawika, L., Roberts, A., Biderman, S., Scao, T. L., & Raffel, C. (2022). Crosslingual generalization through multitask finetuning. *arXiv preprint arXiv:2211.01786*.
Almazrouei, E., Alobeidli, H., Alshamsi, A., Cappelli, A., Cojocaru, R., Debbah, M., & Penedo, G. (2023). *The falcon series of open language models*.
Team, G., Anil, R., Borgeaud, S., Wu, Y., Alayrac, J. B., Yu, J., & Ahn, J. (2023). *Gemini: a family of highly capable multimodal models*.
Guo, D., Zhu, Q., Yang, D., Xie, Z., Dong, K., Zhang, W., & Liang, W. (2024). *DeepSeek-Coder: When the Large Language Model Meets Programming–The Rise of Code Intelligence*.
Wang, D., Raman, N., Sibue, M., Ma, Z., Babkin, P., Kaur, S., & Liu, X. (2023). *DocLLM: A layout-aware generative language model for multimodal document understanding*.
Hwang, A., Li, B., Hou, Z., & Roth, D. (2023). *Large Language Models as Sous Chefs: Revising Recipes with GPT-3*.