

Cooking with ChatGPT and Bard: A Study on Competencies of AI Tools on Recipe Correction, Adaption, Time Management and Presentation

* Ahmet Hakan DEĞERLİ ^a , Nevruz Berna TATLISU ^b 

^a Istanbul Bilgi University, Faculty of Applied Sciences, Department of Gastronomy and Culinary Arts, Istanbul/Türkiye

^b Istanbul Gelisim University, Faculty of Applied Sciences, Department of Gastronomy and Culinary Arts, Istanbul/Türkiye

Article History

Received: 18.09.2023

Accepted: 12.12.2023

Keywords

AI

ChatGPT

Bard

Recipe Generation

Gastronomy

Abstract

With its potential use in many areas including food and beverage sector, artificial intelligence has become one of the most prominent topics recently, partly due to the new AI tools. This study evaluates the competencies of ChatGPT (versions 3.5 and 4) and Bard in relation to food recipes by assigning tasks in five different areas: recipe correction, recipe adaptation, recipe detailing, time management, and presentation. The responses were then analyzed. It was observed that ChatGPT 4 outperformed the other tools in recipe correction, time management and presentation tasks while it gave similar results with ChatGPT 3.5 in recipe adaptation and recipe detailing tasks. Bard performed better than ChatGPT 3.5 in recipe correction but performed worse than both tools in all other tasks. Subsequent discussions highlighted the strengths and limitations of the tools. While these tools' scores may not yet outperform a professional chef in the assigned tasks, they can be alternative and supportive assets in the gastronomy field considering their rapid response rates. Along with the potential use of the tools in tasks such as adapting recipes, managing time, and generating presentation ideas, the ongoing development and interaction of AI tools and related technologies could contribute significantly to the food industry in the future.

Article Type

Research Article

* Corresponding Author

E-mail: ahmet.degerli@bilgi.edu.tr (A. H. Değerli)

DOI: 10.21325/jotags.2023.1312

INTRODUCTION

With recent technological developments, concepts such as artificial intelligence and machine learning have come into use in many areas. Tools developed by OpenAI such as ChatGPT, DALL-E, Codex, Whisper, and CLIP are examples of this in the fields of text generation, image generation, code writing, speech recognition, and computer vision (Kolides et al., 2023).

Artificial intelligence also finds its application areas in many sectors, including food and beverage. It can be said that AI, which has applications such as seed selection, temperature, irrigation, and crop tracking in the sense of growing food products, can also be used in many areas related to the processing, storage, and service of food for food and beverage businesses. These include supporting customers in product selection, tracking equipment sanitation, ensuring customer satisfaction and loyalty, monitoring employees and determining whether necessary procedures are followed, transforming food order applications, revenue tracking, and developing new products/recipes (Kumar et al., 2021).

Generation of food recipes is also among the uses of artificial intelligence (Goel et al., 2022). ChatGPT and Bard are capable of performing various written tasks. As generation of food recipes is also considered a natural language processing task (Goel et al., 2022), these tools can also be expected to generate and transform food recipes. However, even though generation of recipes was mentioned as one of the competencies of ChatGPT (Fusté-Forné & Orea-Giner, 2023) and there are studies in the literature where the competencies of ChatGPT are tested in other subjects (Bang et al., 2023; Niszczoła & Rybicka, 2023), no such study has been encountered for food recipe related competencies of neither ChatGPT nor Bard. Therefore, this study aims to test the competencies of ChatGPT and Bard in various tasks such as detecting and correcting errors, detailing recipe steps, adapting recipes for a specific diet, creating time management plans, and describing presentation in given recipes of meals. These tasks are expected to reveal the tools' knowledge on cooking techniques, ingredients, terminology, special diets and their ability on recipe time management, creating appealing food presentations and to use and convey this knowledge. These are also necessary competencies for planning and running a food service operation. Therefore, this study is expected to reveal the limitations and usability of ChatGPT and Bard as auxiliary tools by professionals and academics working in the related fields.

Literature Review

Brief History of AI

Artificial Intelligence is defined as “a system’s ability to correctly interpret external data, to learn from such data, and to use those learnings to achieve specific goals and tasks through flexible adaptation” (Kaplan & Haenlein, 2019, p. 1). For explaining the brief history of AI, Alan Turing is probably an excellent point to start. Turing developed a machine in order to break the code of Enigma during WW2 which can be considered the forerunner of modern computers (Haenlein & Kaplan, 2019). He was also the first person to mention the intelligent and learning machines scientifically and proposed what is now called “Turing Test” in his article in 1950, a test for determining if a computer is capable of think like a human being (Turing, 1950).

The first use of the word “Artificial Intelligence” in 1956 and the development of some early systems like ELIZA and General Problem Solver in the 1950s and 1960s followed this. These were successful programs of their time but

failed to replicate human intelligence, and thus there has to be a paradigm shift. These early programs tried to replicate human intelligence with collection of rules and a top-down approach, and this approach was able to be successful for tasks like playing chess. But they had to be able to interpret and learn from data their self for many other tasks that are impossible to be expressed by a human being with if-then statements. Although theories that could remedy this problem related with artificial neural networks (ANL) can be traced back to the 1940s, the technology did not allow it. Until the 2010s, with deep learning, a specific type of ANL, today's AI applications became possible (Haenlein & Kaplan, 2019).

AI in Food Industry

Gastronomy 4.0 is a term derived from Industry 4.0 which expresses the 4th industrial revolution after the 1st (mechanization), 2nd (intensive use of electrical energy), 3rd (widespread digitalization) (Lasi et al., 2014) and include the use of technologies such as IoT (Internet of Things), artificial intelligence nanotechnology, big data etc. (Mutlu Öztürk, 2020).

Applications of Gastronomy 4.0 involves Robot chefs and waiters, use of 3D printers, use of augmented reality and digital menus in restaurants in order to enhance the dining experience and Artificial Intelligence (Keskin & Sezen, 2021; Mutlu Öztürk, 2020)

Artificial intelligence has many uses in the food industry including agricultural production, industrial processing, and food and beverage establishments. It's applications can be summarized as the facilitation and automation of processes such as seed selection, soil and crop monitoring, irrigation, and temperature control in agriculture (Kumar et al., 2021) and ensuring food safety, predicting shelf life, estimating quality, cutting and sorting food in the field of industrial food manufacturing thanks to technologies like "Artificial Neural Networks, Fuzzy logic, Computer Vision, and deep learning" (Addanki et al., 2022).

In addition to agricultural and industrial food production, artificial intelligence also has various applications for restaurant businesses. Recommendation engines for meals; robots for food preparation and service; various data management and programming tasks such as taking reservations, food delivery, scheduling staff shifts, inventory tracking, and creating purchasing lists; chatbots; face and fingerprint recognition technologies (Addanki et al., 2022; Dani et al., 2022), monitoring of equipment sanitation; ensuring customer satisfaction and loyalty; monitoring employees and determining whether necessary procedures are followed; and developing new products or generating food recipes are among these current or future applications (Kumar et al., 2021). Some of these applications can be explained like below (Dani et al., 2022; Kızıldemir & Çerkez, 2020; Kumar et al., 2021):

By integrating recommendation engines into digital menus, kiosks, and websites, and having access to customers' previous orders (and ratings), AI can suggest menu items to customers, facilitating their food choices.

Creating purchase lists requires predicting future product sales, calculating the corresponding raw material quantities, and monitoring the inventory of available products. With the help of artificial intelligence, sales data can be used to generate projections, and the necessary material quantities can be calculated based on standard recipes. The products in stock can be entered manually, or computer vision and the Internet of Things (e.g., RFID chips and readers) can be employed to partially or fully automate this process.

The implementation of AI in handling reservations, scheduling staff shifts, and managing food delivery services optimizes operational workflows. Smart reservation systems can analyze historical data to predict peak hours, enabling restaurants to optimize staff schedules and minimize labor costs. AI-powered algorithms can also streamline the delivery process by optimizing routes, predicting delivery times, and enhancing overall logistics efficiency.

AI-powered recipe generation tools can assist chefs in experimenting with innovative combinations for creating new menu items. Furthermore, AI can aid in optimizing menu offerings based on factors such as profitability, ingredient availability, and sustainability. By considering these variables, AI algorithms can suggest menu items that strike a balance between customer appeal and operational feasibility, contributing to the overall success of the restaurant.

Use of AI for Recipe Generation

As food recipes are texts consisting of a list of ingredients, instructions, and sometimes equipment, the generation of recipes by artificial intelligence is considered a natural language processing task under semi-structured text generation. For this purpose, different text generation models and their combinations can be used (Goel et al., 2022). Creating recipes using artificial intelligence may save the trouble of finding a recipe suitable for the ingredients a person intends to use in the meal. Also, it can be used in creating creative recipes (Lee et al., 2020). Applications in this regard can include creating recipes from food photos (reverse cooking), creating personalized recipes according to user preferences and diets, and adapting recipes to different cuisines (Goel et al., 2022; Goel & Bagler, 2022).

The mentioned applications require specific inputs according to their designs and give the recipe as output. For example, Ratatouille (Goel et al., 2022) accepts the desired ingredients as input and gives the name, ingredient, and instruction as output. Chaudhary et al. (2022) developed an application for creating Indian cuisine-centered recipes that similarly takes one main ingredient and other ingredients as input and produces an output consisting of ingredients and instructions. The application developed by Lee et al. (2020) takes the recipe name and ingredients as input and gives the instructions as output or takes the recipe name and instructions as input and gives the ingredients as output.

While these applications use text as input and output, there are other applications that take an image as input or provide an image as output. For instance, in a study conducted by Noever & Noever (2023), a tool was developed that could create a list of products in a refrigerator from a photograph taken and then produce a recipe that could be made with these items. It has been stated that with the method followed in this study, tools could be developed in the future that include features such as taking into account people's dietary restrictions and the preparation time of meals and adjusting portion sizes according to the individual. In the study conducted by Salvador et al. (2019), a tool was developed that predicts the ingredients used in the preparation of a meal from a given food photo, then uses the visual and predicted ingredients to create preparation instructions, thus creating the recipe. It was noted that this tool performed better than other (retrieval-based) approaches.

So far, the applications specifically developed for creating recipes were mostly inaccessible. Ratatouille2 (Goel et al., 2022), although accessible, has encountered explicit errors (nonsensical cooking times, half-finished recipes, illogical instructions) in the trials we made.

ChatGPT and Bard

In addition, tools like ChatGPT and Bard, which were not produced with a specific purpose like creating recipes, can be used for the same purpose and give more consistent results than the mentioned tool. These tools only take text as input, so creating a recipe from images like food or refrigerator is not possible. However, since they allow much more unrestricted inputs and outputs than many tools developed specifically for creating recipes, it is possible to give many tasks such as creating a menu or recipe on a desired theme, improving or creating alternatives on a given recipe, predicting the cuisine/recipe of the given ingredients, etc.

The first of the mentioned tools, ChatGPT, is a natural language processing model based on the Generative Pre-trained Transformer (GPT) model developed by the OpenAI company. ChatGPT 3 version was made available in November 2022, and the ChatGPT 4 version on March 14, 2023 (Rudolph et al., 2023).

The initial studies on the tool focused more on the ethical aspect of ChatGPT (Jabotinsky & Sarel, 2022; Susnjak, 2022), but there are also studies examining the strengths and weaknesses of the tool (Niszczoła & Rybicka, 2023; Bang et al., 2023). In the study conducted by Bang et al. (2023), the competencies of ChatGPT were tested in the context of many tasks based on summarization, translation, question answering, detecting false information, and many tasks based on knowledge or reasoning. It was seen that ChatGPT, although making mistakes, performs better in many tasks than previous models, especially performs well in deductive and abductive logical reasoning tasks, but has limitations in some non-textual reasoning abilities (mathematical, temporal, spatial) and has hallucination problems including producing unreal information, like other models.

In the study conducted by Niszczoła & Rybicka (2023), the reliability of diet lists created by ChatGPT related to food allergy and calorie restriction was examined. As a result of the study, it was seen that in 4 of the 56 tasks given, allergens were not correctly excluded. In none of the diets related to calorie restriction was a nutritional supplement included, and only in one was there a warning about the requested calorie restriction (1000 kcal), and it also made some mistakes regarding calorie and food quantities. On the other hand, it was stated that it has strong aspects such as restricting red meat consumption in favor of fish and white meat consumption, including vegetables/fruits in every meal, providing information about allergens, and reading labels. Although creating meal menus and recipes is also mentioned among the competencies of ChatGPT (Fusté-Forné & Orea-Giner, 2023), no study has been encountered that tests the competencies of ChatGPT in this field.

Previous studies also investigated the potential uses of ChatGPT in tourism industry (Erul & Işın, 2023; Ülkü, 2023). ChatGPT's inability to access information newer than September 2021 (Bal Ram & Pratima Verma, 2023) creates a disadvantage for its practicality in the tourism industry. On the other hand with the use of some ChatGPT 4 plugins it may access to current information on internet and with the use of API (Application Programming Interface) and fine tuning models, it can also be possible to customize GPT based models according to needs and integrate them into the operation (OpenAI, 2023; Ülkü, 2023). Therefore ChatGPT and GPT-based models can have wide uses in the tourism industry such as "facilitating tourists information and decision processes, making reservations, staff training, sustaining operational efficiency, creation of marketing materials, integration with IoT (Internet of Things) systems and overcoming language barrier" (Erul & Işın, 2023; Ülkü, 2023).

Another artificial intelligence tool similar to ChatGPT is Bard, which Google launched on March 21, 2023 (Rudolph et al., 2023). Due to the novelty of the tool, studies conducted on Bard are even more limited. Existing studies generally consider Bard along with other language models and focus on topics such as comparing the tools with each other (Bal Ram & Pratima Verma, 2023; Rahaman et al., 2023; Rudolph et al., 2023) or using these tools in scientific writing (Crawford et al., 2023).

One of the significant differences between Bard and ChatGPT is that while ChatGPT's knowledge is limited to September 2021 and earlier, Bard is a "convolutional" tool (evolving over time on its own) with access to current knowledge (Bal Ram & Pratima Verma, 2023).

In this study, the competence of ChatGPT (3.5 and 4) and Bard tools on tasks to be given related to recipes was tested.

Methodology

Previous studies lists indicators of chefs' competencies and these include "knowledge of food ingredients and technics, time management skills, artistic creativity" along with other indicators such as emotional control, professionalism, motivation and communications skills, budgeting etc. (Birdir & Pearson, 2000; Mahfud et al., 2019; Zopiatis, 2010). As the purpose of this study is to examine the competency of a computer program social aspects of chefs' competency is not included here and former aspects are examined.

The research used five main tasks, namely recipe correction, recipe adaptation, recipe detailing, recipe time management, and presentation. The ability of recognizing and correcting mistakes regarding ingredients and technics (recipe correction), the ability of substituting ingredients with suitable ones (recipe adaptation) and the ability of detailing a recipe by explaining technical terms for ingredients and techniques (recipe detailing) reflects "knowledge of food ingredients and technics". Creating and expressing food presentations reflects "artistic creativity" and creating time management plans for given recipes reflects "time management skills"

Short explanations related to the tasks are given in Table 1. More detailed explanations of the tasks are given under the related sub-headings. Each task was given separately to ChatGPT's 3.5 and 4 versions and Bard. Since the relevant tools perform their best in English, using them in this language was preferred. Tasks are completed between 13th of June and 5th of July.

Table 1. Tasks Given to the Tools

	Test	Preliminary Information	Task	Used Recipes
1	Recipe correction	Ingredients, preparation steps	List the errors in terms of ingredients and techniques	-Hollandaise sauce -Chocolate chip cookie
2	Recipe detailing	Ingredients, preparation steps	Make the recipe, written briefly and with technical terms, more detailed and understandable.	-Osso buco
3	Recipe adaptation	Ingredients, preparation steps	Adapt the meal for an individual following a special diet such as vegan, keto, diabetic, etc.	-Osso buco -Chocolate chip cookie -Hollandaise sauce
4	Recipe time management	Ingredients, preparation steps	Plan time management for the steps of recipes planned to be made simultaneously	-Osso buco -Chocolate chip cookie -Hollandaise sauce -Beef Wellington*
5	Presentation	Meal name	Describe an aesthetic/appetizing presentation for the specified meal and create an SVG code representing the presentation.	-Beef Wellington -Osso buco*

*Used in additional task

Since the tools do not produce the same answer each time, especially in tasks 1 and 3, where quantitative evaluation is made, the reliability of the results has been tried to be increased by producing answers to the same prompt repeatedly or repeating the same task with different recipes.

Recipe Correction

For this task, some ingredients and stages in the used recipes were changed to cause problems in the final product, making the recipes faulty. Then, tools were prompted “Can you tell me what's wrong with the following recipe?” following the modified recipes. The tasks were repeated five times for each recipe, and the errors detected by the tools each time were determined. The changes made in the recipes are given in Table 2.

Table 2. Modifications in the Recipes

Hollandaise		Chocolate Chip Cookie	
Original	Modified	Original	Modified
Water should not touch the bowl	Water should touch the bowl	Baking temperature 175 C	Baking temperature 150 C
Be careful not to overheat the eggs	Make sure the eggs are well warmed up	Cookies are placed on the tray 5 cm apart	Cookies are placed on the tray 1 cm apart
Cover and store in a warm place	Refrigerate the sauce and serve cold	After the cookies are cooled on the tray for a while, they are taken to the cooling wire	Cookies are placed on the cooling wire immediately after baking
½ Cup Butter	2 Cups Butter	3 Cups Flour	6 Cups of Flour

Recipe Detailing

A brief osso buco recipe with an instruction part with 51 words including technical terms such as "dusting, searing, sauteing, pinching, deglazing, braising, bouquet garni and mirepoix" was provided to tools following the prompt “Can you turn the following recipe into a more understandable manner”

Recipe Adaptation

The original versions of the Osso Buco, Chocolate Chip Cookie, and Hollandaise sauce recipes were given following the prompt “Can you adapt the following recipe for ‘vegan, vegetarian, keto, gluten-free’ diets separately, specifying the amounts?” The ingredients in the original recipes and the substitutions each tool made for each diet were examined for correctness.

Time Management

Recipes were provided to tools following the prompt “I will make the following recipes, can you tell me how should I manage the time?” Using osso buco, chocolate chip cookie and hollandaise sauce firstly and then adding beef wellington along with these for the additional task with the same prompt. Then, created time managements was examined in regard to calculating times correctly, making use of passive times, etc.

Presentation

Tools were prompted with "Can you describe an attractive presentation for Beef Wellington and create an SVG code to represent the presentation?" in order to obtain a depiction and a SVG image regarding the presentations of the food.

Findings

Recipe Correction

For the hollandaise sauce, ChatGPT 3.5 reached a score of 9/20, while ChatGPT 4 reached 17/20. While ChatGPT 3.5 and 4 listed potential errors in the recipe as a list of items as an answer, Bard both listed potential errors and created a corrected recipe. On the other hand, it was seen that Bard corrected more errors in the edited recipes than it listed, but it could not manage to list all these errors. Accordingly, when the edited recipes are also considered (marked with y), Bard detected the errors the same number of times (17) as ChatGPT 4. However, if only the errors it listed are considered (marked with x), it was able to detect the errors only five times.

Table 3. Hollandaise Recipe Correction

Hollandaise Recipe Correction	ChatGPT 3.5					ChatGPT 4					Bard				
	Regeneration					Regeneration					Regeneration				
	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5
Bain-marie			x				x	x	x		y	y	y	y	y
Heating the eggs	x		x		x	x	x	x	x	x		y	y	x	x
Service Temperature			x	x	x	x	x	x	x	x			x	x	x
Butter Quantity	x		x				x	x	x	x	y	y	y	y	y
Total Score:	9/20					17/20					17/20 (5/20)				

For the Chocolate Chip Cookie recipe, ChatGPT 3.5 reached a score of 11.5, while ChatGPT 4 scored 15 out of 20. ChatGPT 3.5 suggested using 4-5 cups of flour incorrectly in one instance, which was a problem related to the quantity, while ChatGPT 4 consistently corrected this to 2-3 cups of flour. On the other hand, ChatGPT 3.5 noticed three times that the necessary cooling step on the tray, which is crucial to harden the cookies, was skipped, while ChatGPT 4 never noticed.

Table 4. Chocolate Chip Cookie Correction

Chocolate Chip Cookie Correction	ChatGPT 3.5					ChatGPT 4					Bard				
	Regeneration					Regeneration					Regeneration				
	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5
Oven Temperature	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
Distance Between Cookies		x	x	x		x	x	x	x	x					
Cooling		x	x	x							x	y	y	x	y
Flour Quantity			?			x	x	x	x	x	x	x	x	y	x
Total Score:	11.5/20					15/20					15/20 (11/20)				

Similarly, Bard corrected the same number of errors (15) as ChatGPT 4 in the modified recipe but did not manage to list all of them. Furthermore, Bard suggested in every instance of the revised recipes that the cookies should be left on the tray before moving them to the cooling rack. However, it was only able to express on the fourth attempt that "leaving them on the tray for a while will help them maintain their shape". In the second and third instances, it incorrectly stated that it was essential to move the cookies to the cooling rack immediately after baking, claiming that "otherwise, they will stick to the rack" (2nd regeneration) and "they will continue to bake and dry out on the tray" (3rd regeneration).

Recipe Detailing

Both versions of ChatGPT provided instructions averaging 300 words, sticking to the steps in the recipe and detailing them correctly, while ChatGPT 4 also added the labels chopped/minced next to vegetables and started the instructions section with the cutting of vegetables.

On the other hand, Bard provided a set of instructions of 168 words and a "tips" section of 118 words. While generally detailing the recipe correctly, it made certain changes like merging the steps of adding tomatoes and bouquet garni (removing the 10-minute cooking period in between), changing the reduction ratio of the wine (from 1/3 to 1/2), and therefore not sticking to the original recipe as ChatGPT.

Recipe Adaptation

We were able to categorize the changes in the adapted recipes under six groups, namely "correct substitutions, incorrect substitutions, unnecessary substitutions, unnecessary omissions, being extra-cautious and uncalled revisions".

Table 5. Substitution Behaviors and Counts

Substitution Behaviour	ChatGPT 4	ChatGPT 3.5	Bard
Right substitutions	13	12	8
Wrong substitutions	-	-	2
Unnecessary substitutions	5	2	9
Unnecessary omissions	1	1	8
Being extra-cautious	4	3	2
Uncalled revisions	4	2	11

Correct substitutions refer to replacing non-diet-friendly ingredients with appropriate ones. Incorrect substitutions refer to replacing an ingredient with another that is unsuitable for the diet or unavailable. Unnecessary substitutions refer to replacing diet-friendly ingredients with another diet-friendly ingredient. Unnecessary removals refer to removing diet-friendly or non-diet-friendly ingredients from the recipe instead of keeping or substituting them. Unwanted corrections refer to unrelated (unit change, using stock instead of water, using dry instead of fresh ingredients, etc.) corrections. Some examples encountered are given in the table below. Accordingly, no clear difference was seen between ChatGPT 3.5 and 4, while Bard performed significantly worse.

Table 6. Substitution Behaviors

Type	Ingredient	Vegan	Vegetarian	Keto	Gluten-Free
Right Substitutions	Osso Buco	Portobello Mushrooms	Portobello Mushrooms	-	-
	All-Purpose Flour	-	-	Almond Flour or Coconut Flour	Gluten-Free All-Purpose Flour
Wrong Substitutions	Osso Buco	Seitan Osso Bucos	Boneless, Skinless Chicken Thighs	-	-
Unnecessary Substitutions	All-Purpose Flour	Almond Flour Or Gluten-Free Flour Blend	Whole Wheat Flour	4 Boneless, Skinless Lamb Shanks	4 Boneless, Skinless Chicken Thighs

Table 6. Substitution Behaviors (cont.)

Being Extra-Cautious	Lemon	-	-	Omit If Desired	
	Chocolate Chips	Vegan Chocolate Chips	Dark Chocolate Chips (Ensure It's Vegetarian-Friendly)	-	Gluten-Free Semisweet Chocolate Chips
Uncalled Revisions	Water	Vegetable Broth	Vegetable Broth	Beef Broth	Beef Broth
	Pinch Of...	1/4 Teaspoon	1/4 Teaspoon	1/8 Teaspoon	1/4 Teaspoon

Time Management

All three tools generally sequenced the recipes as osso buco, chocolate chip cookie, and hollandaise sauce. ChatGPT 3.5 briefly summarized these three recipes and the steps to be performed in order, while ChatGPT 4 divided the recipe steps into three headings: "preparation, cooking, completion and service". The first step involves preparing the mise en places for all the recipes and making the cookie dough, the second step involves cooking the osso buco, cookies, and sauce in this order, and the final step involves plating.

On the other hand, Bard briefly summarized the recipes in the same order and then provided a timeline. Here, starting the cookie dough between sealing and braising the osso buco (despite no passive time), and incorrectly naming the hollandaise sauce as lemon curd were noticeable. Each tool followed a similar order and predicted a total time of about three hours. Considering that Osso buco is the product with the longest cooking time among the given products, the cooking time provides enough time for the preparation of the other two products, and considering that the chocolate chip cookie is served cold, the order followed by ChatGPT 3.5 seems to be the most efficient plan to produce the three products in the shortest time. On the other hand, the plan provided by ChatGPT 4, which starts the plan with mise en place and is divided into three parts as preparation, cooking, and serving, might be more efficient to be applied in a professional kitchen.

The same task was repeated by adding another recipe with more active and passive time steps (Beef Wellington). Initially, Bard only provided general tips on time management in cooking and cooking times for the given recipes. When asked for a detailed plan, it provided a timetable where the cookies and Hollandaise sauce were prepared in the first stages and osso buco, cookies and beef wellington is baked at 2:15, 2:30 and 2:45 respectively, with all items finished by 3:00(!).

ChatGPT 3.5 fundamentally followed the sequence of osso buco, hollandaise, Beef Wellington, chocolate chip cookie, while ChatGPT 4 followed the sequence of osso buco, Beef Wellington, chocolate chip cookie, and hollandaise. While ChatGPT 3.5 generally kept the recipes in blocks and could not utilize the passive times except for the cooking times of osso buco and Beef Wellington, ChatGPT 4 carried out other preparations during the passive times in the Beef Wellington recipe and left the hollandaise sauce for the last to serve it fresh. As a result of this task, ChatGPT 4 suggested the most successful time management.

Presentation

It was seen that ChatGPT versions produced detailed depictions including the position, size, color, and shape of the materials to be placed. Despite being repeated several times, Bard could not understand the given prompt, provided answers such as giving a food recipe, describing a powerpoint presentation, or listing general suggestions about the presentation, and could not create a viewable SVG code. When the prompt was updated as "Can you

describe an attractive Beef Wellington plate by specifying the shape, location, angle, etc. of each material on the plate?" Bard still described a less detailed presentation than the two versions of ChatGPT and generated a less detailed image in response to the prompt "Can you create a SVG code representing the presentation you described?"

Table 7. Beef Wellington Presentation Descriptions

<p>ChatGPT 3.5</p>	<p>“1. Plating: - Start by placing a square or rectangular plate in the center. - Position a small bed of vibrant green baby spinach leaves at the center of the plate. - Carefully place the cooked Beef Wellington on top of the spinach, allowing it to stand upright. 2. Garnishes: - Surround the Beef Wellington with a drizzle of rich, velvety mushroom sauce. Pour it generously over the top and let it flow onto the plate. - Add a few sprigs of fresh herbs, such as thyme or rosemary, around the dish for an aromatic touch. - Sprinkle some finely chopped parsley or chives on top of the Beef Wellington to add a pop of color. 3. Sides: - Accompany the dish with complementary sides like roasted baby potatoes or garlic mashed potatoes, arranged neatly beside the Beef Wellington. - Add a small portion of sautéed seasonal vegetables, such as carrots, asparagus, or green beans, to provide freshness and balance to the plate.”</p>
<p>ChatGPT 4</p>	<p>“Imagine a large, rectangular white porcelain plate. At the center, resting on a thin layer of rich, dark wine reduction is the golden-brown Beef Wellington, its flaky pastry crisp and gleaming. To the side of the Beef Wellington is a small pile of lightly sautéed baby spinach, its vibrant green creating a lovely contrast with the deep hues of the meat and sauce. At the other end of the plate, there is a small ramekin of the same wine reduction for those who prefer extra sauce. Thinly sliced and fanned out roasted beets add a splash of color to the plate. Sprinkles of freshly chopped parsley are scattered across the dish, adding a final touch of freshness.”</p>
<p>Bard</p>	<p>“The beef Wellington is the centerpiece of the plate, fanned out slightly so that the golden brown crust is visible. It is placed in the center of a large, oval platter. The side dish, such as roasted vegetables or mashed potatoes, is arranged in a mound to the side of the beef Wellington. A sprig of rosemary is placed on the side of the platter, and a drizzle of sauce is added to the beef Wellington”</p>

The images created by the tools are given in Figures 1, 2, and 3. Since the visuals created by the ChatGPT versions do not include the names of the ingredients, the ingredient names have been added to the visuals in the way they are specified in the SVG codes for clarity. Since there are no names related to the elements in Bard's SVG code other than those added to the visual by itself, no such process has been performed.

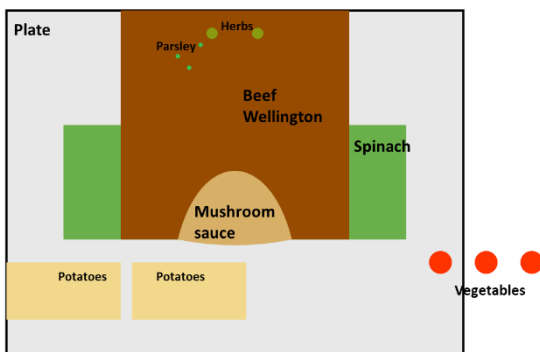


Figure 1. Visual Created by ChatGPT 3.5

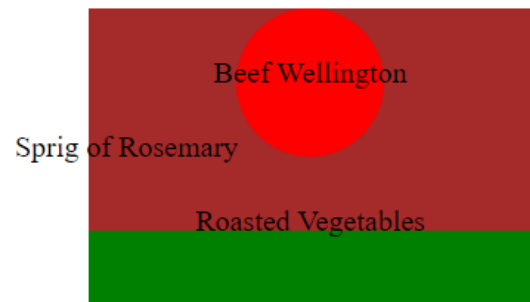


Figure 2. Visual Created by Bard

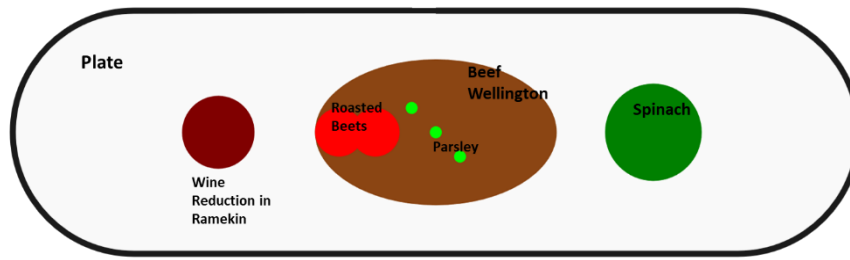


Figure 3. Visual Created by ChatGPT 4

When the descriptions and visuals created were examined, it was seen that Bard creates much less detailed descriptions and visuals compared to the ChatGPT versions. While ChatGPT 3.5 includes the plate in the visuals it creates, and sticks to the descriptions and product colors it has made, it has been observed that ChatGPT 4 additionally sticks to the sizes of the products, it places all the products inside the plate and pays attention to leaving negative space. The same situation was observed when the same task was repeated with another dish (Figures 4 and 5).

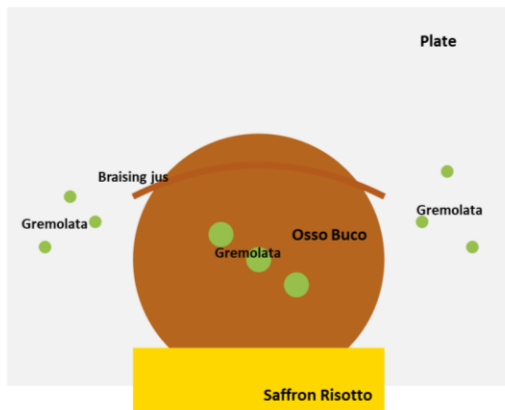


Figure 3. Visual Created by ChatGPT 3.5 (Osso Buco)

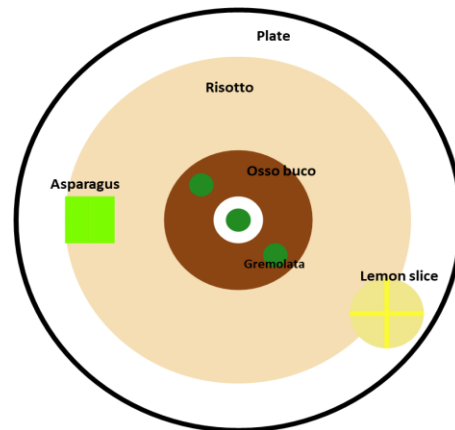


Figure 4. Visual Created by ChatGPT 4 (Osso Buco)

Discussion and Conclusion

In this study, ChatGPT 3.5, ChatGPT 4, and Bard models were given 5 different tasks related to food recipes, and the competencies of the tools in these areas were tested. In 3 of the given tasks (1, 4, and 5), ChatGPT 4 performed noticeably better, and in 2 (2 and 3), it performed similarly to ChatGPT 3.5.

In the recipe correction task, the most successful tool was ChatGPT 4, with a performance accuracy of 80%. While ChatGPT 3.5 performed with about 50% accuracy, Bard's performance can be evaluated on a wide scale ranging from 30% to 80% due to inconsistencies in the responses it generated as it failed to point out mistakes but was able to provide corrected recipes.

Both ChatGPT versions completed the recipe detailing task successfully. Despite generally demonstrating a successful performance, Bard made some unsolicited changes in the recipe which seems problematic.

In the recipe adaptation task, both ChatGPT versions performed well and adapted the given recipes so that no ingredients violated the requested diet. Bard failed to demonstrate the same performance and made more unsolicited changes compared to other models, similar to the previous task. While Niszczota & Rybicka, 2023 found that

ChatGPT may create diet menus that violate dietary restrictions (food allergies), neither version of ChatGPT has shown any violations for the diets (vegan, vegetarian, gluten-free, and keto), unlike Bard.

In the time management task, ChatGPT 4 clearly performed the best, and Bard performed the worst. While ChatGPT 3.5 managed to create a feasible plan by correctly evaluating the preparation times of the products but it could not perform as well as ChatGPT 4 in evaluating passive times and serving products fresh.

In the task of preparing a presentation, unlike the ChatGPT versions, Bard could not design a pleasant presentation and describe it in detail, nor could create a SVG code for a detailed image. While there is no significant difference in terms of description among ChatGPT versions, ChatGPT 4 has been noticeably more successful in visualizing their presentations with SVG code.

Bard's responses were more inconsistent and unsolicited comparing to ChatGPT versions as it provided two incompatible list of mistakes and corrected recipe as an answer to recipe correction task; made unasked changes and omissions in recipe detailing task; and made highest number of unnecessary substitutions, unnecessary omissions and uncalled revisions in recipe adaptation tasks

The use of the tools was generally quite simple. When written in a clear language, especially the ChatGPT models almost always managed to understand the tasks correctly. However, there were instances where the Bard model misunderstood the task or performed unsolicited actions.

Although the release of artificial intelligence models like ChatGPT, Bard, Dall-E, Mid-Journey to the public has been the subject of the last few years, the performances demonstrated by ChatGPT 4 in particular in the given tasks are impressive. While it may not match the level of expertise demonstrated by a human chef or dietitian in terms of quality across most tasks, it significantly outperforms an individual with no training or experience in these fields. Its remarkable achievements in tasks like providing detailed recipes and managing time, coupled with the fact that these models generate responses almost instantly, make a strong case for acknowledging their capabilities.

Of course, there are some predictable differences between a human chef and a natural language processing model. For example, a chef has the opportunity to actually cook and observe parameters such as applicability, taste, texture while improving a recipe or adapting it for a diet, and make corrections accordingly. The performances demonstrated by ChatGPT or Bard are acquired only from text sources in the data they were trained on. Therefore, they have some limitations due to being developed solely based on texts. For instance, as shown in the study by Bang et al. (2023), while it can provide a very successful and detailed description when asked to describe a flag, it produces incorrect results for flags containing relatively complex shapes like leaves, stars when asked to create an SVG code for the same flag's drawing. Bang et al. (2023) stated that this situation is a flaw of text-only language models and it's almost impossible to draw a leaf without seeing it in reality. In the same study, it was also found that ChatGPT's spatial reasoning ability is lower than many other types of reasoning.

For an entity that has never seen any image and never been in space, these restrictions are not surprising at all. Similarly, not having smelled and tasted will significantly reduce a chef's abilities.

On the other hand, considering technologies like electronic nose, robot chefs, the performance of AI tools, which exhibits these performances with their current training, could reach beyond the limits of human imagination in the future with access to data like aroma molecules and texture data related to foods, feedbacks from human chefs and

robot chefs, changes occurring in foods as a result of different cooking and preparation techniques, responses given by individuals from different societies to these foods etc.

Chatelan et al., 2023 stated that with its current state AI tools can assist nutrition and dietetics professionals but cannot replace them, these study has a similar outcome. Although these tools do not currently demonstrate a performance that will surpass a good chef or dietitian, when their response times are taken into account, they seem to be potential auxiliary tools that will facilitate the work of professionals in the food sector, as in many other sectors, and increase their performance. As also shown in this study, these tools can be used for troubleshooting recipes, adapting recipes for different needs, creating time management charts, and generating new presentation ideas, and also they may be used for other tasks such as creating menus, editing menu texts, organizing employee shifts, etc. Of course as there may be mistakes as the results of these tasks always a person should check the final results, but still use of the AI tools may help for easing these tasks and coming up with creative ideas.

Limitations

The study was conducted between June 13 and July 5, 2023; therefore, the results do not encompass potential improvements that may have been made to the tools after this period. During the specified dates, ChatGPT-4 had not yet activated the feature of receiving prompts with visual inputs, and since such a feature was not present in other tools neither, no tasks involving visual inputs were utilized. In future research, the performance of ChatGPT in visual tasks can be evaluated. Although it provides an idea of their performance in similar tasks, the competence of the tools has only been tested in five tasks. In future studies, competencies in other tasks such as menu planning, creating staff shifts, generating order lists, etc., can be assessed.

Declaration of Generative AI and AI-assisted technologies in the writing process

During the preparation of this work the author(s) used ChatGPT 4 in order to edit and translate the text. Furthermore author(s) used Bard and ChatGPT for the implementation of methods due to the nature of the study. After using these tools/services, the author(s) reviewed and edited the content as needed and take(s) full responsibility for the content of the publication.

Declaration

All authors of the article contributed equally to the article process. The authors have no conflicts of interest to declare.

REFERENCES

- Addanki, M., Patra, P., & Kandra, P. (2022). Recent advances and applications of artificial intelligence and related technologies in the food industry. *Applied Food Research*, 2(2), 100126. <https://doi.org/10.1016/J.AFRES.2022.100126>
- Bal Ram, & Pratima Verma. (2023). Artificial intelligence AI-based Chatbot study of ChatGPT, Google AI Bard and Baidu AI. *World Journal of Advanced Engineering Technology and Sciences*, 8(1), 258–261. <https://doi.org/10.30574/wjaets.2023.8.1.0045>

- Bang, Y., Cahyawijaya, S., Lee, N., Dai, W., Su, D., Wilie, B., Lovenia, H., Ji, Z., Yu, T., Chung, W., Do, Q. V., Xu, Y., & Fung, P. (2023). A Multitask, Multilingual, Multimodal Evaluation of ChatGPT on Reasoning, Hallucination, and Interactivity. *ArXiv Preprint ArXiv:2302.04023*. <http://arxiv.org/abs/2302.04023>
- Birdir, K., & Pearson, T. E. (2000). Research chefs' competencies: a Delphi approach. 12(3), 205–209. <http://www.emerald-library.com>
- Chatelan, A., Clerc, A., & Fonta, P. A. (2023). ChatGPT and Future Artificial Intelligence Chatbots: What may be the Influence on Credentialed Nutrition and Dietetics Practitioners? *Journal of the Academy of Nutrition and Dietetics*, 123(11), 1525–1531. <https://doi.org/10.1016/j.jand.2023.08.001>
- Chaudhary, S., Soni, B., Sindhavad, A., Mamaniya, A., Dalvi, A., & Siddavatam, I. (2022). ChefAI.IN: Generating Indian Recipes with AI Algorithm. 2022 *International Conference on Trends in Quantum Computing and Emerging Business Technologies, TQCEBT 2022*. <https://doi.org/10.1109/TQCEBT54229.2022.10041463>
- Crawford, J., Cowling, M., Ashton-Hay, S., Kelder, J. A., Middleton, R., & Wilson, G. S. (2023). Artificial Intelligence and Authorship Editor Policy: ChatGPT, Bard Bing AI, and beyond. *Journal of University Teaching and Learning Practice*, 20(5). <https://doi.org/10.53761/1.20.5.01>
- Dani, R., Singh Rawal, Y., & Bagchi, P. (2022). *Opportunities and challenges in implementation of artificial intelligence in food & beverage service industry*. <https://doi.org/10.1063/5.0103741>
- Erul, E., & Işın, A. (2023). ChatGPT ile Sohbetler: Turizmde ChatGPT nin Önemi (Chats with ChatGPT: Importance of ChatGPT in Tourism). *Journal of Tourism and Gastronomy Studies*, 11(1), 780–793. <https://doi.org/10.21325/jotags.2023.1217>
- Fusté-Forné, F., & Orea-Giner, A. (2023). Gastronomy in tourism management and marketing: an interview with ChatGPT. *ROBONOMICS: The Journal of the Automated Economy*. <https://www.zangador.institute>
- Goel, M., & Bagler, G. (2022). Computational gastronomy: A data science approach to food. In *Journal of Biosciences* (Vol. 47, Issue 1). Springer. <https://doi.org/10.1007/s12038-021-00248-1>
- Goel, M., Chakraborty, P., Ponnaganti, V., Khan, M., Tatipamala, S., Saini, A., & Bagler, G. (2022). Ratatouille: A tool for Novel Recipe Generation. *Proceedings - 2022 IEEE 38th International Conference on Data Engineering Workshops, ICDEW 2022*, 107–110. <https://doi.org/10.1109/ICDEW55742.2022.00022>
- Haenlein, M., & Kaplan, A. (2019). A Brief History of Artificial Intelligence: On the Past, Present, and Future of Artificial Intelligence. *California Management Review*, 61(4), 5–14. <https://doi.org/10.1177/0008125619864925>
- Kaplan, A., & Haenlein, M. (2019). Siri, Siri, in my hand: Who's the fairest in the land? On the interpretations, illustrations, and implications of artificial intelligence. *Business Horizons*, 62(1), 15–25. <https://doi.org/10.1016/J.BUSHOR.2018.08.004>
- Keskin, E., & Sezen, N. (2021). Gastronomi 4.0 Üzerine Kavramsal Bir Araştırma. *Gastroia: Journal of Gastronomy and Travel Research*, 5(2), 177–198.
- Kızıldemir, Ö., & Çerkez, M. (2020). Yiyecek-İçecek İşletmelerinde Yapay Zekâ Kullanımı. *Türk Turizm Arastirmalari Dergisi*, 4(2), 1264–1278. <https://doi.org/10.26677/tr1010.2020.394>

- Kolides, A., Nawaz, A., Rathor, A., Beeman, D., Hashmi, M., Fatima, S., Berdik, D., Al-Ayyoub, M., & Jararweh, Y. (2023). Artificial intelligence foundation and pre-trained models: Fundamentals, applications, opportunities, and social impacts. *Simulation Modelling Practice and Theory*, 126. <https://doi.org/10.1016/j.simpat.2023.102754>
- Kumar, I., Rawat, J., Mohd, N., & Husain, S. (2021). Opportunities of Artificial Intelligence and Machine Learning in the Food Industry. *Journal of Food Quality*, 2021. <https://doi.org/10.1155/2021/4535567>
- Lasi, H., Fettke, P., Kemper, H. G., Feld, T., & Hoffmann, M. (2014). Industry 4.0. *Business and Information Systems Engineering*, 6(4), 239–242. <https://doi.org/10.1007/s12599-014-0334-4>
- Lee, H. H., Shu, K., Achananuparp, P., Kokoh Prasetyo, P., Liu, Y., Lim, E.-P., & Varshney, L. R. (2020). RecipeGPT: Generative Pre-training Based Cooking Recipe Generation and Evaluation System. 17. <https://doi.org/10.1145/nnnnnnn.nnnnnnn>
- Mahfud, T., Pardjono, & Lastariwati, B. (2019). Chef's competency as a key element in food tourism success: A literature review. *Geojournal of Tourism and Geosites*, 26(3), 1057–1071. <https://doi.org/10.30892/gtg.26329-417>
- Mutlu Öztürk, H. (2020). Teknolojik Gelişmeler ve Gastronomi Alanına Yansımaları: Gastronomi 4.0. *Güncel Turizm Araştırmaları Dergisi*, 4(2), 222–239. <https://doi.org/10.32572/guntad.703872>
- Niszczoła, P., & Rybicka, I. (2023). The credibility of dietary advice formulated by ChatGPT: Robo-diets for people with food allergies. *Nutrition*, 112, 112076. <https://doi.org/10.1016/j.nut.2023.112076>
- OpenAI. (2023). <https://openai.com/>
- Rahaman, Md. S., Ahsan, M. M. T., Anjum, N., Rahman, Md. M., & Rahman, M. N. (2023). The AI Race is on! Google's Bard and Openai's Chatgpt Head to Head: An Opinion Article. *SSRN Electronic Journal*. <https://doi.org/10.2139/ssrn.4351785>
- Rudolph, J., Tan, S., & Tan, S. (2023). War of the chatbots: Bard, Bing Chat, ChatGPT, Ernie and beyond. The new AI gold rush and its impact on higher education. *Journal of Applied Learning and Teaching*, 6(1), 364–389. <https://doi.org/10.37074/jalt.2023.6.1.23>
- Turing, A. M. (1950). Computing Machinery and Intelligence. Source: *Mind, New Series*, 59(236), 433–460. <http://www.jstor.orgStableURL:http://www.jstor.org/stable/2251299> Accessed:25/08/200818:56
- Ülkü, A. (2023). ChatGPT-4 for Hospitality: Implications. *Journal of Tourism and Gastronomy Studies*, 11(3), 1727–1743. <https://doi.org/10.21325/jotags.2023.1263>
- Zopiatis, A. (2010). Is it art or science? Chef's competencies for success. *International Journal of Hospitality Management*, 29(3), 459–467. <https://doi.org/10.1016/j.ijhm.2009.12.003>