Food Media Content Study for an AI Smart Speaker

Kyoung-Ah Kim[1*]

Abstract

Society advances through technology, and technology has changed many lifestyles. The need for food is varying, but the availability of food is constantly changing as trends in production change. Combining the food industry and technology, a robot that delivers food and also cooks it has been developed. The time has come for a combination of food content and technology to advance the restaurant industry. This study discusses the application of a recommended food content media providing system using a curation engine that recommends contents according to individual tastes and preferences for the convenience of those who use food contents, using artificial intelligence speakers. We discuss the technologies required to develop video contents optimized for AI speakers with screens and shapes, combined with inset top boxes.

Key Words: AI smart speaker, Food media content, Recommendation system, Video content tagging information.

I. INTRODUCTION

The AI speaker market is growing rapidly and is attracting attention in the ICT and media industries. After Amazon Echo was developed, Samsung, Google, and Microsoft also entered the AI speaker market, developing a variety of AI or voice-controlled speakers, and in 2018 Amazon and Google also launched AI speakers with displays. Voice applications are gaining attention in the AI smart speaker market. Apple, which developed Siri, and Amazon, which developed Alexa, are collaborating with many developers to showcase various voice applications. Applications that reflect the needs of consumers using artificial intelligence speakers are changing the market.   The smart speaker market in Korea is being developed using mobile carriers and portals such as KT, SKT, Naver, and Kakao. The technology has rapidly spread through the only form of sale in the world, a combination of artificial intelligence speakers and set-top boxes. However, due to the lack of diversity of service contents and the unsatisfactory applications available, the utilization rate of AI for food platforms is significantly lower in Korea than in other countries. In addition to set-top boxes, AI speakers have been combined with kitchen appliances such as refrigerators, ovens, and ranges, and increasing numbers of

people are using smart cooking appliances. The importance of developing a food content platform that can provide speed and convenience is evident.

A voice recognition speaker based on an AI platform can penetrate various objects and spaces that a display-oriented interface cannot utilize. Such a speaker is therefore is advantageous for multitasking. The food content provided by existing platforms does not consider the incorporation of smart speakers, but simply broadcasts video or makes simple recommendations such as today's dish or the platform's own recommendation. Therefore, these approaches cannot convey the value of necessity or convenience by utilizing the food content platform of users of the AI speaker. In this study we investigated the technology necessary to service a food content media platform using AI speakers. These technologies are not limited to food applications, but can be used in various application fields

In this paper, we discuss video production technology and service direction effective to increase the utilization rate of food contents, providing convenience for AI speaker users, and to provide various food contents production and services accordingly. First, we describe a system for providing recommendations for food contents, using a curation engine. The curation engine makes it easy to search for customized recipe information according to a user's

tastes and preferences, provide information about appropriate ingredients, and link to order delivery sites where recipe materials can be purchased. The second technique described relates to the design of video content information section mapping. This function allows inputting of food-related information into each time period of a video. Users can leave feedback on each section and collect user profiles through data shared on social media, to increase the reliability of the curation engine. It is possible to purchase ingredients by entering the product information of the ingredients in each section. The two technologies discussed in this paper are essential elements for providing contents to users more efficiently for all generations who search for and collect all information by video. Food content created by applying these technologies will change the time and space in which people consume their diets. This will increase the usage rate of food contents for AI speaker users and inspire new and diverse food contents production. Therefore, this paper proposes a way to realize a food media contents service optimized for AI speaker.

II. RELATED RESEARCH

As online information about food, recipes, and ingredients grows exponentially, users can spend considerable time finding the information they need. Efficient online curators have emerged to speed up the search for specific information on the Internet. Chef Watson, developed by IBM in 2014, analyzes hundreds of existing recipes and helps users to create new recipes. It learns recipes by identifying recipe templates, generating new ingredient combinations by matching chemical components, assessing ingredient combinations, and creating new recipe steps derived from existing templates
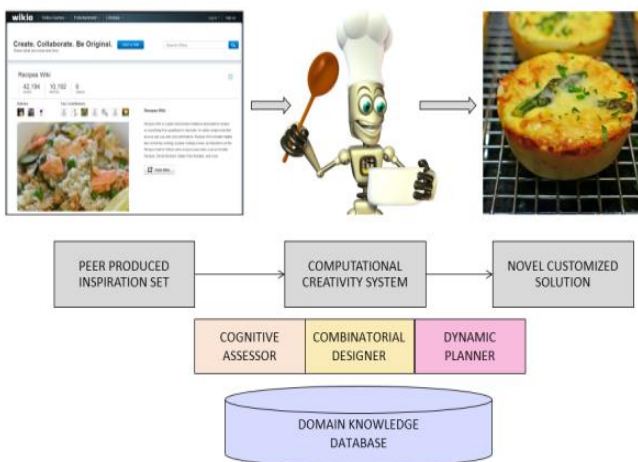

Fig. 1. IBM Chef Watson cognitive cooking.

Domestically available food content curation platforms include recipe recommendations, restaurant recommendations, and delivery apps. These platforms provide content by inferring user preference through self-recommendation or by collecting user information.

The cooking recommendation system and method shown in Figure 2 includes several steps: (A) providing a Web page including an interface for inputting cooking ingredients; (B) retrieving a recipe associated with the ingredients and generating a first search result, extracting ingredients other than the ingredients input into the first webpage from the retrieved recipes; (C) generating
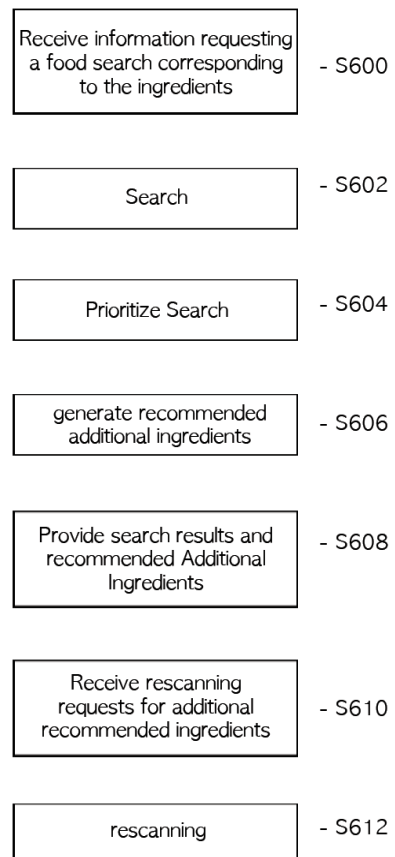

Fig. 2. Cooking Recommendation System and Methods.

additional recommended ingredient information; (D) providing a second web page including the first search result generated from the searched document and additional recommended ingredients, additional recommended ingredients selected from the additional recommended ingredients; and (E) generating a second search result by searching for a recipe associated with the ingredients input to the first Web page and the selected additional recommended ingredients. However, since this system is limited to recommending recipes by searching for recipes according to the information input, there is a limitation to the provision of recipe information optimized for the needs

of users according to their emotions, environment, or situation. In addition, there are uncertainties such as those associated with user intentions and emotions, sensors, and causality in the platform environment and the user's disposition. To overcome this uncertainty, it is necessary to develop a platform based on video in order to effectively analyze user profiles for recommended contents. Recently, Amazon and Google have launched AI screen-mounted speakers, and Naver, Kakao and KT are about to launch their versions of the speakers.

In order to provide food image contents optimized for AI-equipped display speakers, user information is collected and segmented through identification of video food contents that infer tastes and preferences through specific feedback and sharing by users. This study adopted a curation system that recommends ingredients by applying a food specialization classification system, an ontology, which includes factors such as locality, culture, taste, weather, season, and sensitivity.

III. RECOMMENDED FOOD CONTENT MEDIA PROVISION SYSTEM

The recommended food content media providing system using the curation engine can solve technical problems that can provide recipe information optimized for the various needs of users according to their emotions, environment, or situation, in addition to the availability of ingredients. The recommended food content media providing system using the curation engine is a food DB that stores text data about food information, pictures of ingredients and cooking tools, pictures of recipes step by step, and cooking videos as content profiles, and which processes the attribute information (Figure 3).

The system includes a classification unit for categorizing food information such as material, place, region, culture, taste, weather, season, situation, atmosphere, and emotion; a query input unit for receiving queries about food as text input, a menu selection or an interactive type; a curation engine for analyzing query contents, extracting keywords and related keywords, analyzing similarity between food information from the content profile, and analyzing the taste and preference of the user profile to recommend optimized food information from the classification unit; and a food content providing unit configured to selectively display and display the ingredients corresponding to the recommended food information by combining text about a recipe, a picture for each cooking
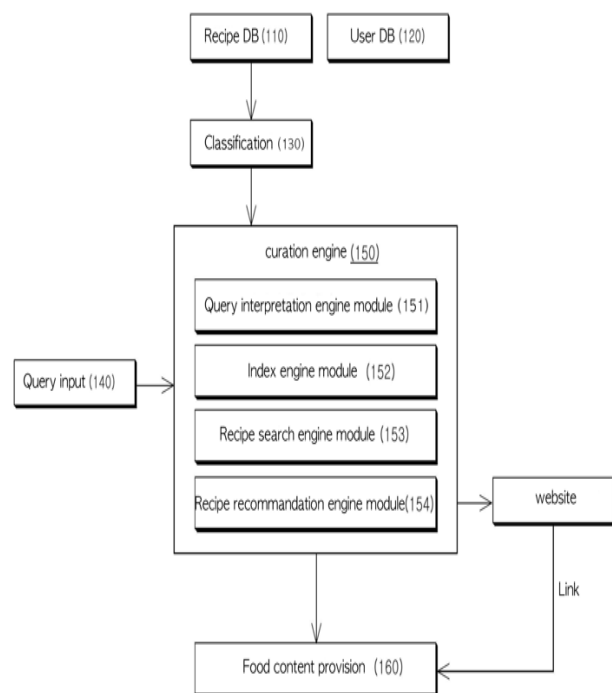

Fig. 3. Recommended food contents media provision system.

step, or a cooking process video. The curation engine unit may recommend food according to its similarity to food in use. The classification unit categorizes food information into food and ingredients, food genre related to weather or season, emotions related to food, health information related to food ingredients, and dishes recommended for the environment and situation. It can be categorized according to the curation history of food information. In addition, the curation engine may link to food material order sites to purchase the relevant ingredients and products needed for the selected recipe. The text information includes cooking instructions, cooking time, amounts of ingredients, basic information about the cooking process, the characteristics and flavors of each ingredient, and details about the origin and possible substitutes for the ingredients. It may include information related to the situation, the atmosphere, the weather, and accompanying beverages.

The curation engine unit converts and extracts query contents into text form, and uses a query module analysis engine module for morphological analysis of the extracted query contents, an index engine module for indexing and storing analysis results through morphological analysis according to similarity It consists of a food search engine module that expands a search term through the analysis result through the

analysis and searches for food information from the index using extended search terms, and a food recommendation engine module which creates recommendations based on the retrieved food information. From the results of a query it is possible to recommend relevant food information with high similarity by assigning a recommendation weight to a specific morpheme corresponding to a particular food, material, weather, season, environment, place, or emotion. The food content update server may periodically update the information of the food content DB through a communication network, or search for and download food information independent of the query contents.

The curation engine makes it easy to search for customized food information according to a user's taste and preference using text input, selection from a menu or interactively. It can provide corresponding food contents, and tag the information in each section of the food image contents. Frequency analysis provides customized food image information or related product information tailored to the user. In addition, the ability to link to an ordering site from which the ingredients can be purchased adds to the convenience of the system.

IV. FOOD CONTENT MEDIA PROVISION SYSTEM IMPLEMENTATION EXAMPLE

In this study, we develop a platform applying a food content media providing system to an AI speaker service. The service contents are divided into five levels. In the first task, the user asks questions according to the individual's context, such as cooking, ingredients, situation, weather, season, taste, and emotional state. The AI speaker analyzes and answers these questions and select additional questions that further elucidate the user's preference. The second task is to recommend dishes or recipes that suit your taste. In the third task, the user can select the type of food content service among text, image, and video. This allows transmission to any device connected to an AI speaker, ensuring space-time autonomy.

The fourth task is to check the content selected by the user. Fifth, the user can go to the connected purchase site through the content provided by the platform to purchase related materials or products. The description is based on recipe contents for understanding the implementation example of the food contents media providing system. However, through the provision of various genres of contents combined with food, the purpose of using the contents is not just for users who intend to eat. It will also
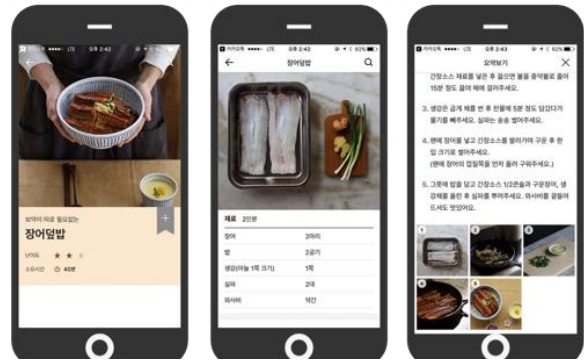

Fig. 4. Food contents media platform implementation example.

contribute to the new food culture and the food service industry by creating content that can be practical and appealing to all users of AI speakers. In order to implement this system, research on the system of collecting user profiles and the food specialization system [material, region, culture, taste, weather, season, situation, emotion] should be done.

V. VIDEO SEGMENT MAPPING BASED ON VIDEO CONTENT TAGGING INFORMARION

The video content information section tagging service collects the tags and section information from the videos watched by the viewers on the service platform, and then analyzes the relationships between users and tags using natural language analysis algorithms and provides customized video and related information to users based on the curation information extracted.
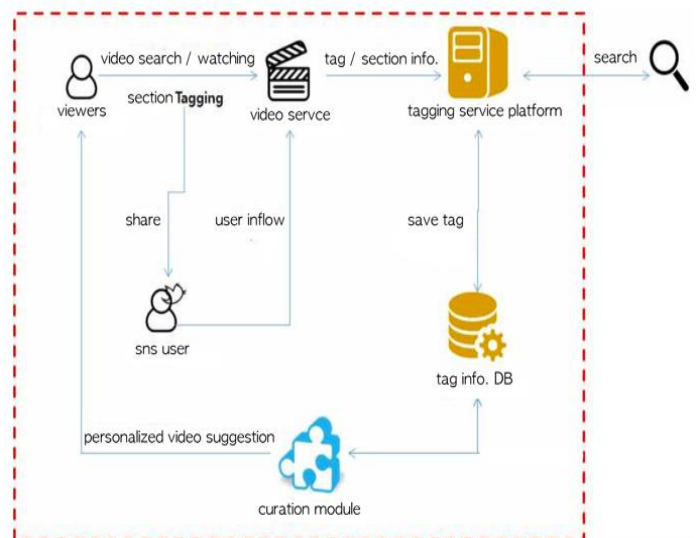

Fig 5. Video tagging service process.

The "Development of Tagging Information Based Image

Segment Mapping and Natural Language Analysis Technology for Customized Video Contents Recommendation Service", a project which was researched and developed as part of an R & D project of the Ministry of Culture, Sports and Tourism, demonstrated that the use of artificial intelligence speakers combining video content tagging technology with effective food content platform services is possible.

The video content that can be provided on the platform are not limited to recipe videos but include video contents such as cooking education videos and cooking shows with entertainment elements that can induce the user's interest are also key tags for information on various ingredients and products. Ingredients and products can be entered and saved as metadata. By combining the food content media providing system using the above-mentioned curation engine, it is possible to produce food content optimized for AI speakers equipped with displays. The user can adjust the video content to be played at a desired time through voice recognition using key tag input to each section. This function can increase the user's use of food image content through the artificial intelligence speaker.

The key tag information entered by the user when uploading each video is registered in the system through the filtering process, in order to prevent error operations caused by inscription and word usage, and the filtered key tag is also registered in the user's video.
The key tags of each video may also be related to the video through user comments in the key tag input by the first registrant.
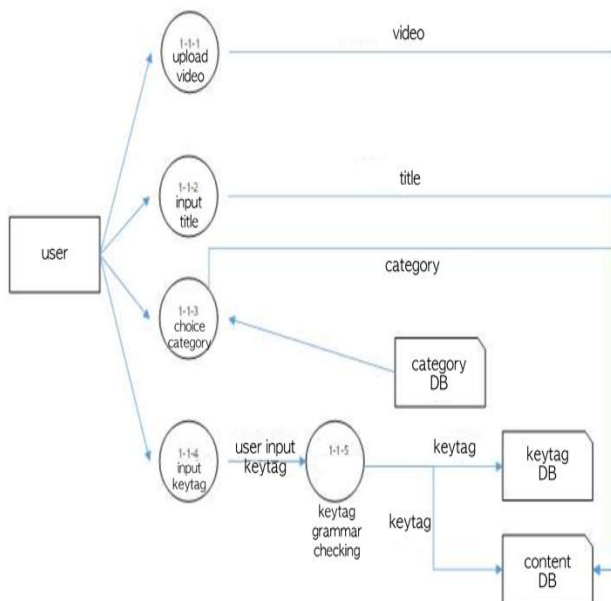

Fig. 6. Tagging information content registration/modification.

Use sentence case for the title. Do not use capitals for

author's surname. Add "and" before the last au-thor. Do not add a period after the last keyword. Keywords are extracted using natural language processing from user tag information and video subtitles, voice recognition, and video recognition through the video URL database. We are building a database based on a relational database using the mSTUV platform. The data consist of a video ID and a timeline as identifiers, and the tagging includes automatic tagging and user comments. The video is given meta information from the analysis of tagging of the video, and data mining based on video and audio information. This meta information is continuously updated and recommended to users. The data can be viewed as the basic information of the system.
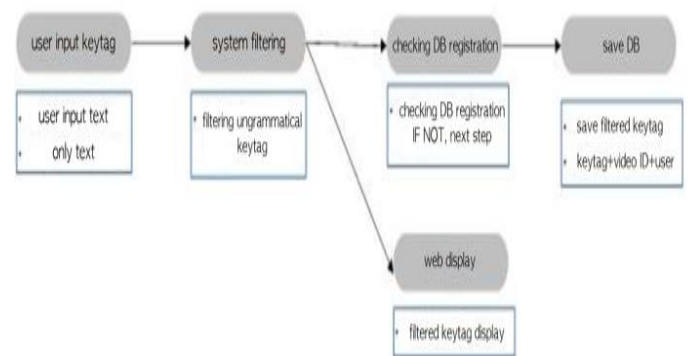

Fig. 7. Definition of the key tag input method.

With the research and development of video segment mapping technology based on video content tagging information, AppLab added a curation engine function and combines existing the tagging function and search function through the mSTUV platform. Based on user information, past search and viewing information, it has been presented business models in travel, film, and cosmetics.
Users who use food content with tagging information-based video segment mapping technology through AI speakers can obtain information about everything in the video in real time using voice recognition, in addition to just watching the desired video. A user can obtain the right content and products, and sellers can find the right consumer.

VI. CONCLUSION

Food has characteristics that are enjoyed through all human senses, and for this reason, artificial audio only has a limit in provision of food contents and increasing the utilization of users. Therefore, it is necessary to produce and provide food contents with video. In this study, we discussed the method of applying the curation engine and

image content information mapping technology to apply a food classification system to service food contents through AI speaker. In addition to information on the contents input by the information provider, users can quickly and easily select desired information by enabling curation through the user profile from the tag inputs of the user. Video content based on key tagging information, video content section mapping technology, and curation engine-based media providing system using this food image content are all content services or educational service platforms in the catering industry that can be serviced by AI speakers equipped with displays. This process is necessary for the advancement of domestic AI speaker diffusion, diversity of food content production, and the development of appropriate technology, and above all, for the convenience and effective provision of content for users using the AI speaker.

The combination of food content and technology will play a role in improving the food service market, breaking down the barriers of time and space for individuals' food culture, and lifestyle, and enhancing content consumption and food culture value.

REFERENCES

[1] Jae-wook Ahn, "Cognitive Cooking with Chef Chef Watson", in *Proceeding of Conference of Korean Society of Food Science and Nutrition*, pp. 89-89, 2016.

[2] Florian Pinel and Lav R. Varshney, "Method forDynamic Ingredient Substitution in Recipes", Application No.13/972,232, https://researcher.watson.ibm.com/researcher/view_group.php?id=5077, 2013.

[3] Korean Registered Patent Publication No. 0850569 (Cooking Recommendation System and Method. 2008.07.30).

[4] M-Lab Co., Ltd., Korea Contents Agency, "Video segment mapping and natural language analysis technology based on tagging information for customized video content recommendation service" (Research report R2017050033/Result of R & D of 2017 Cultural Technology R & D Free Project Project) p. 5, 2017.

Author

Kyoung-Ah Kim received his BA degree in the Department of French Language and Culture from KwanDong University Korea, in 2006 and MA degrees in the Department of Gastronomy from Boston University, USA, in 2016. In 2018, she joined the Department of Cultural Content for pursuing her Ph.D. degree at Dongguk University.

As a producer and director, she has enjoyed highly successful career at Food TV in Korean and acquired the knowledge about food in USA. She is a CEO of Potlucktable Co., Ltd. (food content production), work for education, create food contents and food culture.

Her research interests include uniting Food media content and technology for AI smart speaker.