

# e-Government portal for providing integrated agrifood information based on Big Data Analysis

Euijoo Yeo<sup>1</sup>, Sunggoo Yoo<sup>2</sup>, Yunhwa Koh<sup>3</sup>, Junghoon Moon<sup>4</sup>

## INFO

Received  
Accepted  
Available on-line  
Responsible Editor:

### Keywords:

e-Government, big data  
analysis, agrifood information

## ABSTRACT

The e-government portal is a service that provides integrated agrifood information. Its purpose is to integrate vast amounts of agrifood data and to provide user-customized information on agrifood. Public agrifood data were collected from the Korean government, while private agrifood data were collected from Social Network Services and blogs. We analyzed the data and classified consumer purchasing behavior into five types. We then analyzed agrifood preference by type and weather conditions. Based on the results of these analyses, we developed recommendations for appropriate menus and agrifood and provided price information of recommended agrifood to consumers.

## 1. Introduction

“The age of Big Data has begun.” Recently, vast amounts of data have been generated from various sources, such as the Internet, cloud systems, and mobile devices (Dumbill 2013). Due to technological advances, the data can be gathered and analyzed through computers and algorithms. In this context, discussions on how to organize and use this enormous amount of data have continued in the academic world and in practical areas (Waller and Fawcett 2013).

In Korea, the Ministry of Agriculture, Food, and Rural Affairs (MAFRA) and government-affiliated organizations have accumulated vast amount of agrifood data. MAFRA, for example, has collected agrifood data related to agrifood businesses, and the National Agricultural Products Quality Management Service has collected agrifood certification data. Furthermore, various organizations, including firms and nonprofit organizations, have accumulated agrifood data. Even though the agrifood data is valuable for farmers, agribusinessmen, and consumers, these data have been scattered in many different databases and there has been no integration so that the data are not utilized effectively. In this light, MAFRA has identified the need to integrate the agrifood data, and thus the e-Government portal has been developed for this purpose.

The aim of this study is to introduce the recently developed e-Government portal in the area of agrifood in Korea and present the process of portal development. The services of the e-Government portal include 1) providing integrated agrifood data and 2) providing user-customized information on agrifood.

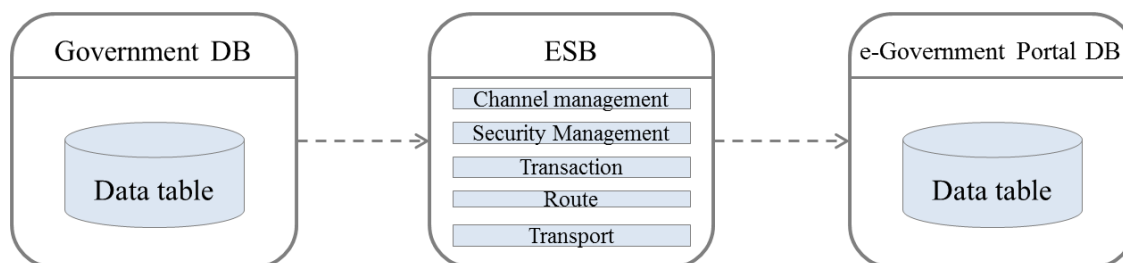
## 2. Methods

### 2.1. Data collection

Agrifood data for the e-Government portal were collected from government organizations, Korean households, and the private sector. Public agrifood data were collected from MAFRA and government-

<sup>1</sup> Euijoo Yeo, Sunggoo Yoo, Yunhwa Koh, Junghoon Moon (corresponding author)  
Department of regional information, Seoul National University, Seoul Korea  
[eiyeo@snu.ac.kr](mailto:eiyeo@snu.ac.kr), [Kaiseven@snu.ac.kr](mailto:Kaiseven@snu.ac.kr), [yunhwakoh@gmail.com](mailto:yunhwakoh@gmail.com), [moonj@snu.ac.kr](mailto:moonj@snu.ac.kr)

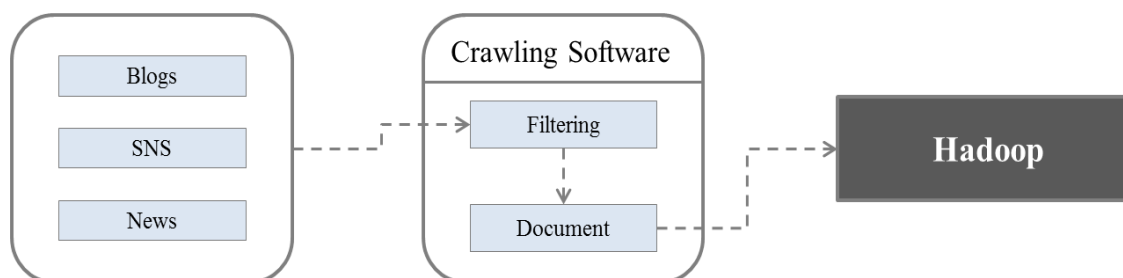
affiliated organizations through a database. The database of each organization was connected to the e-Government portal using the Enterprise Service Bus (ESB) software, which integrates a service-oriented architecture and provides new services (Schmidt et al. 2005). Currently, it is widely used by the Korean government. Through this process, the agrifood public data were collected automatically and saved in the e-Government portal database. Figure 1 illustrates the process of collecting the public agrifood data.



**Figure 1.** Process of collecting public agrifood data

Consumers' purchasing agrifood data were collected from 640 Korean households. The data include the actual agrifood purchasing data and the data on consumers' food-related lifestyles (FRLs). The actual purchasing data were obtained from Korean households' food purchase receipts, which include information on the name of the product, the amount purchased, the date, and the price. The data of consumers' FRLs were collected through surveys. The FRL is measured by five items: method of shopping, quality aspects of food, cooking method, consumption situation, and purchase motive (Grunert et al. 2011).

Private agrifood data were collected from blogs, online shopping malls, news articles, and Social Network Services (SNS) using the crawling technique. Raw agrifood data from private sectors were collected by crawling software. The crawling software, Talkro CIMS developed by Daum Soft, enables filtering of unnecessary data and effective collection of data on crucial key words of user interest. The collected agrifood data were saved in Hadoop, which is a system that saves vast amounts of data dispersively and effectively. Figure 2 illustrates the process of collecting the private agrifood data.



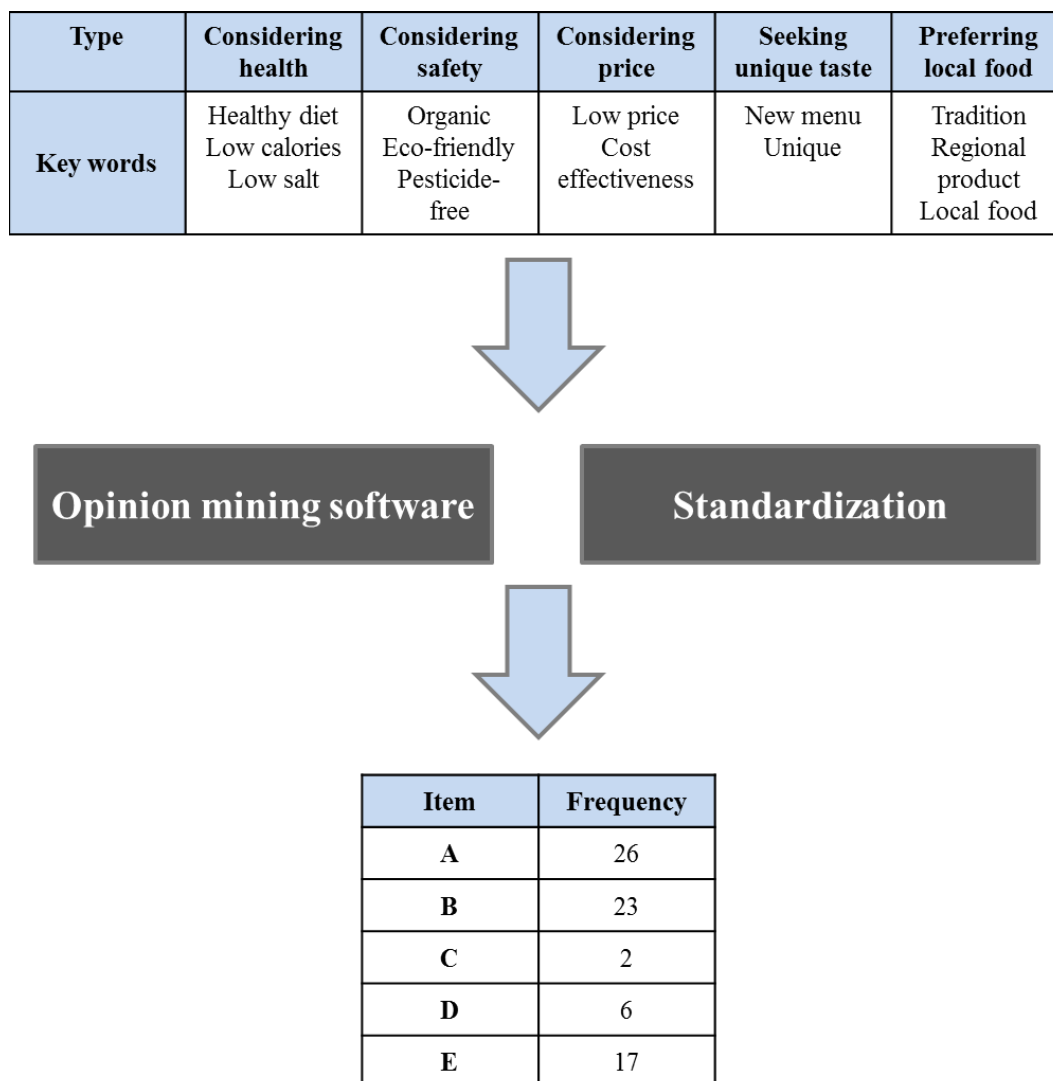
**Figure 2.** Process of collecting private agrifood data

## 2.2. Data analysis

The collected agrifood data were organized and analyzed for developing the e-Government portal. First, consumer purchasing behavior was classified by using cluster analysis. We used factor analysis and extracted 10 factors from the data of consumers' FRLs. By analyzing these factors, we found five significant factors that we defined as consumer's purchasing behavior. These were: considering health, considering safety, considering price, seeking unique taste, and preferring local food.

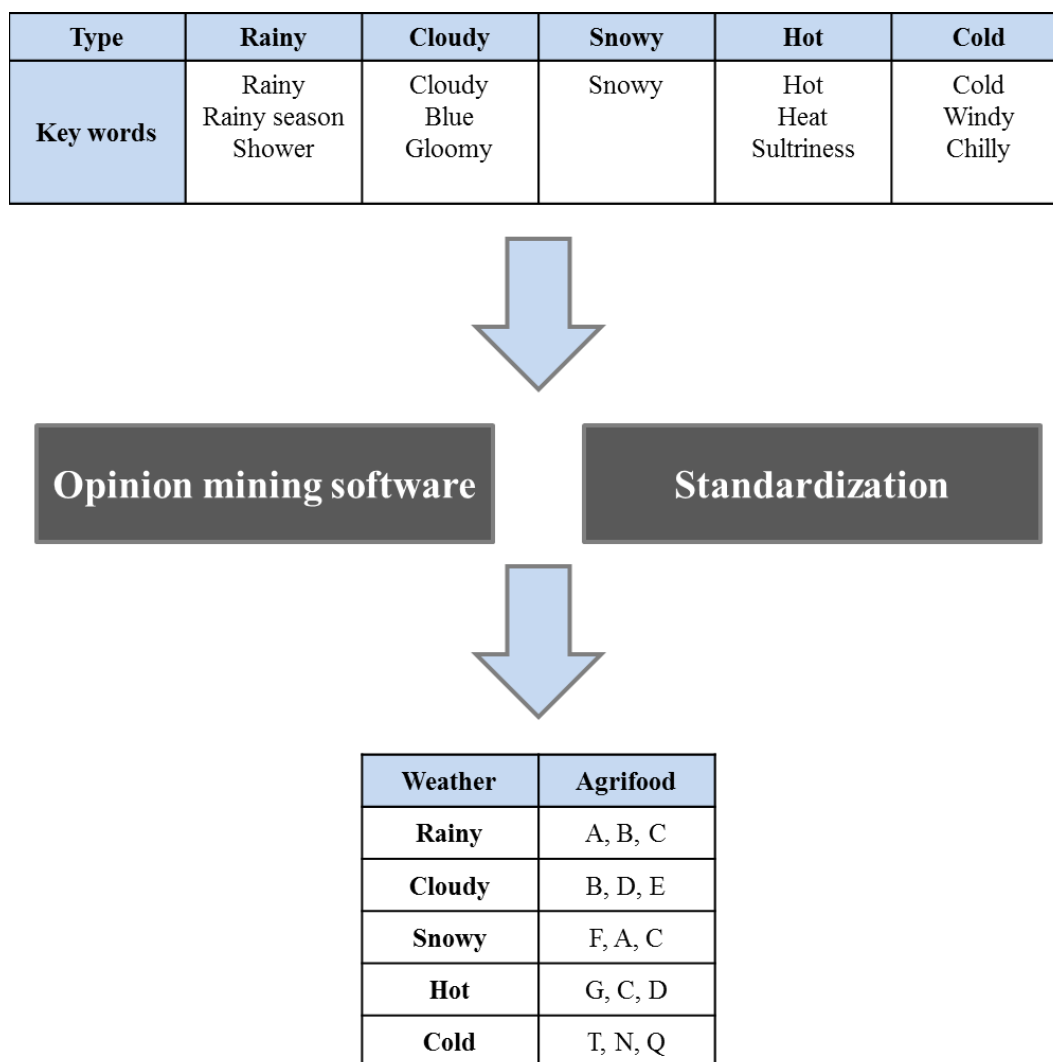
Second, a consumer's preference for agrifood by the five types of purchasing behavior was obtained by analyzing the private agrifood data and the consumer's actual purchasing data. From the actual purchasing data, we listed and then standardized the purchased agrifood items. We then investigated the high-frequency items by each type of purchasing behavior. The private agrifood data were analyzed by using the opinion-mining technique using the opinion-mining software Social Metrics. In this technique,

we input the key words for each type of purchasing behavior and found the related data. In a similar way, we investigated high-frequency items by each type. Figure 3 presents the process of analyzing preferences for agrifood.



**Figure 3.** Process of analyzing preferences for agrifood

Lastly, the information of preference for agrifood with respect to weather and seasonality was obtained by analyzing consumer purchasing data and climate data collected from the Korea Meteorological Administration. The reason why we collected the weather data is that recommendation systems considering seasonality and weather is more preferred by users (Lee and Lee 2006). The process was the same as above and is shown in Figure 4.



**Figure 4.** Process of analyzing preference for agrifood in regard to weather

### 3. e-Government Portal

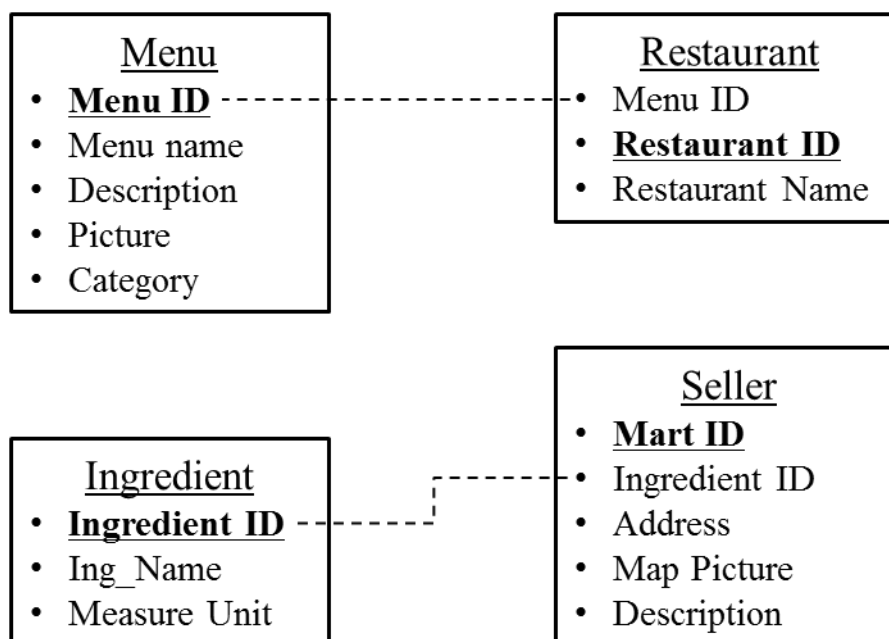
#### 3.1. Organizing the public agrifood information

The collected public agrifood data were organized and effectively presented on the e-Government portal. The scattered public agrifood data were collected by the process described in Section 2.1. However, the database was unstructured and a user-friendly organization was needed. To address this, the agrifood data were arranged according to three standards—types, themes, organizations—and the data were provided in various forms, such as document files, grids, and links, so that users could easily find and identify the appropriate data.

#### 3.2. Consumer recommendation systems

In order to develop recommendation system, we refer to previous research about food-recommendation system. In general, the purpose of food-recommendation system is to provide healthy diet list (Phanich, Pholkul and Phimoltares 2010) and to increase consumer's satisfaction by helping the decision making process of diet choice (Ueta, Iwakami and Ito 2011). Because we identified the type of consumer, the e-Government portal can accomplish the both purpose.

The e-Government portal provides recommendations on the appropriate agrifood for each type of consumer. Consumers can get information on a recommended menu and agrifood, restaurant, and market. Figure 5 illustrates an entity relationship diagram of a recommendation system.



**Figure 5.** Entity relationship diagram of the recommendation system

For more specificity, consumers can obtain information on their type of purchasing behavior through a brief survey. Based on the information, consumers can receive a recommendation of the appropriate agrifood by type. For example, consumers who are concerned with health can get a menu and agrifood recommendation along with detailed information for health. The recommended menu and agrifood have information on price change, seasonality, and preference for agrifood for particular weather conditions. Consumers can then know which agrifood is cheap and appropriate for the current weather.

In the case of a menu, the system provides information on ingredients and restaurants for a recommended menu. The restaurant information includes a list of the top 10 restaurants and frequency and content of mentions on SNS.

In the case of agrifood, the recommendation system provides daily online and offline price information for each agrifood. Furthermore, information on retail and wholesale agrifood markets is indicated, including information on price fluctuations, the price floor, and the location of markets. Figures 6 and 7 illustrate the recommendation system on the e-Government portal.

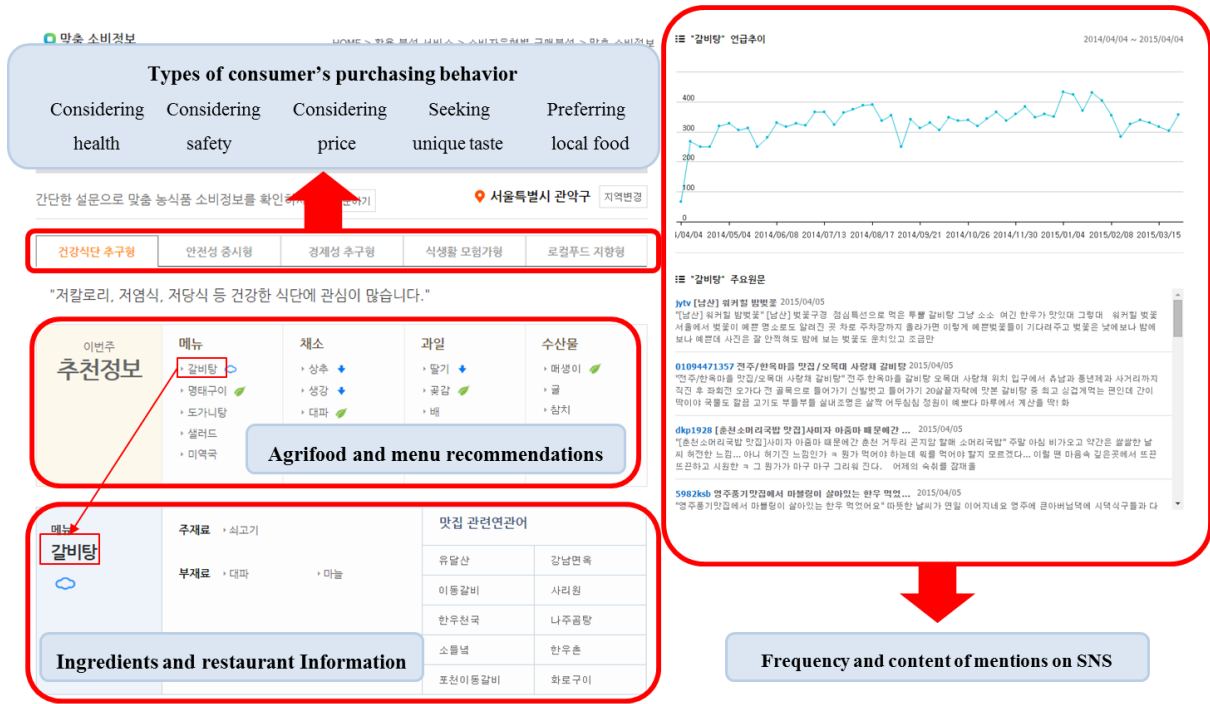


Figure 6. Recommendation system on the e-Government portal

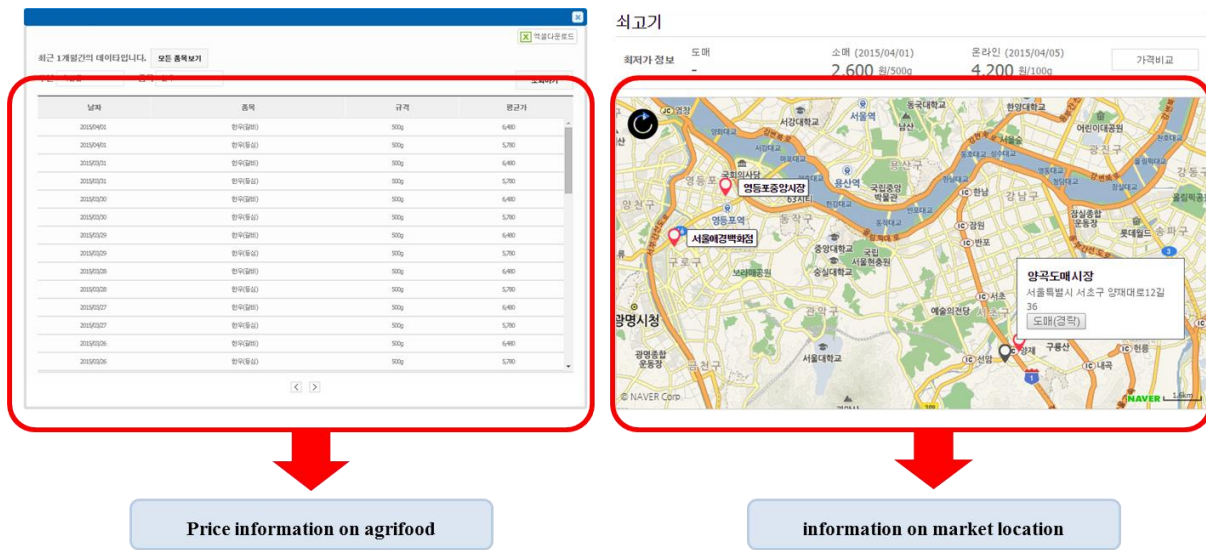


Figure 7. Recommendation system on the e-Government portal

#### 4. Conclusion

In discussing the implications of the e-Government portal, we determined that the e-Government portal creates social value. First, it provides appropriate information regarding agrifood so that consumers can make a rational purchasing decision; this may result in improving consumers' health. Second, the e-Government portal can provide farmers and producers with new business opportunities by offering vast amounts of organized data. Last, the price fluctuation of agrifood can be stabilized because the price of agrifood is declining. Because consumers can easily notice the low price of agrifood, demand for the product will increase, which may help stabilize price fluctuations.

## Acknowledgements

This work was carried out with the Support of “Cooperative Research Program for Agriculture Science and Technology Development (Project No. PJ011390012015)” Rural Development Administration, Republic of Korea.

## References

- Dumbill, E 2013, 'Making sense of big data', *Big Data*, vol. 1, no. 1, pp. 1-2. doi: [10.1089/big.2012.1503](https://doi.org/10.1089/big.2012.1503)
- Grunert, KG, Perrea, T, Zhou, Y, Huang, G, Sørensen, BT and Krystallis, A 2011, 'Is food-related lifestyle (FRL) able to reveal food consumption patterns in non-Western cultural environments? Its adaptation and application in urban China', *Appetite*, vol. 56, pp. 357-367. doi: [10.1016/j.appet.2010.12.020](https://doi.org/10.1016/j.appet.2010.12.020)
- Lee, JS and Lee, JC 2006, 'A case based music recommendation system using context-awareness', *J. KIIS*, vol 12, no. 3, pp. 111-126.
- Phanich, M, Pholkul, P and Phimoltares, S 2010 'Food recommendation system using clustering analysis for diabetic patients', *IEEE.Int. Conf. Inf. Sci. and Appl*, pp.1-8. doi: [10.1109/icisa.2010.5480416](https://doi.org/10.1109/icisa.2010.5480416)
- Schmidt, M-T, Hutchison, B, Lambros, P and Phippen, R 2005, 'The enterprise service bus: making service-oriented architecture real', *IBM Systems Journal*, vol. 44, no. 4, pp. 781-797. doi: [10.1147/sj.444.0781](https://doi.org/10.1147/sj.444.0781)
- Ueta, T, Iwakami, M and Ito, T 2011 'Implementation of a goal-oriented recipe recommendation system providing nutrition information', *IEEE Int. Conf. TAAI*, pp. 183-188. doi: [10.1109/taai.2011.39](https://doi.org/10.1109/taai.2011.39)
- Waller, MA and Fawcett, SE 2013, 'Data science, predictive analytics, and big data: a revolution that will transform supply chain design and management', *Journal of Business Logistics*, vol. 34, no. 2, pp. 77-84. doi: [10.1111/jbl.12010](https://doi.org/10.1111/jbl.12010)