



# INTERNATIONAL JOURNAL ON INFORMATICS VISUALIZATION

journal homepage : [www.joiv.org/index.php/joiv](http://www.joiv.org/index.php/joiv)



## Design of a Big-data-Based Decision Support System for Rational Cultural Policy Establishment

Youngseok Lee <sup>a</sup>, Gimoon Cho <sup>b</sup>, Jungwon Cho <sup>c,\*</sup>

<sup>a</sup> Department of Computer Education, Seoul National University of Education, 96 Sechojungang-ro, Secho-gu, Seoul, Republic of Korea

<sup>b</sup> Department of Industrial Data Science, Kangnam University, 40 Gangnam-ro Giheung-gu, Yongin-si, Gyeonggi-do, Republic of Korea

<sup>c</sup> Department of Computer Education, Jeju National University, 102 Jejudaehakno, Jeju-si, Jeju-do, Republic of Korea

Corresponding author: \*jwcho@jejunu.ac.kr

**Abstract**— This paper proposes a technique for designing a decision-making system based on big data to support rational cultural policy decisions. To identify a rational cultural policy, it is necessary to extract a comparable index for cultural policy and analyze and process factors in terms of cultural supply and cultural consumption. Analyzed and processed supply indices and consumption indices become the basic input data for calculating additional cultural indices that can be measured at the cultural level of each region. Regional cultural indices are treated as independent variables in terms of cultural supply, and target variables are considered in terms of cultural demand. Two corresponding types of regression models are established. Based on the eXtreme gradient boosting and light gradient boosting machine algorithms, which are representative algorithms for calculating cultural indicators, we attempted to construct and analyze a model of the proposed system. The developed model is designed to predict the demand index according to the regional cultural supply index. It was confirmed that the demand side could be changed based on supply-side items by using the proposed technique to support decision-making. Due to the complexity of the policy environment of modern society, mixing various policy tools targeting multiple functions is accepted as a common basis for policy design, but institutional arrangements are needed to reflect the results of various data analyses in budget decision-making. This will be possible to produce data based on effectiveness and suggest appropriate rational policies and decisions.

**Keywords**— Culture supply policy; cultural figures; balanced culture; big data; decision-making system.

Manuscript received 23 Oct. 2021; revised 20 Dec. 2021; accepted 21 Jan. 2022. Date of publication 31 May 2022.  
IJASEIT is licensed under a Creative Commons Attribution-Share Alike 4.0 International License.



### I. INTRODUCTION

The average lifespans of humans are increasing based on the development of modern medical technology. According to a report by the National Statistical Office, the average life expectancy had already exceeded 80 years in 2008. Since 2017, Korea has become an aging society, and 14% of the total population is aged 65 years or older [1]. In such an era, people have become more interested in living “a valuable life,” rather than simply solving food, clothing, and shelter concerns, and are making efforts such as enjoying cultural activities to live more valuable lives. According to the cultural nostalgia survey by the Ministry of Culture, Sports, and Tourism in 2018, the cultural and art viewing rate reached 80% [2].

Additionally, with the implementation of the 52h workweek, the government is making significant efforts to encourage cultural life by implementing cultural policies such

as the Cultural Diversity Policy and Culture Vision 2030 [3]. Many efforts are being made to revitalize culture at the national and local social levels, and cultural supply policies such as constructing cultural facilities and running various cultural events are being implemented to further these efforts [4]. However, most efforts are currently implemented based on the subjective sense of policy executors without objective decision-making methods. In some cases, policy implementation succeeds when the person in charge feels positive, but in other cases, it can lead to major failures [5]. Regarding Korea's data utilization policy establishment, it is necessary to review and analyze the current status of data-based policy establishment, and it is desirable to promote objective policies based on data utilization policies [6]. However, because simple data evidence does not guarantee the effectiveness of a policy, controversy regarding its effectiveness can arise [7]. Therefore, the development direction of data utilization policy is to create guidelines and manuals on data utilization so that data-based decision-

making receives greater attention. It is recommended that the advantages and disadvantages of a system be analyzed and leveraged in stages.

In this paper, we first examine the concept of the cultural index and understand the principles and methods of data-driven decision-making to make decisions based on the cultural index. Moreover, the related research trends are analyzed after examining the principles and processes of prediction techniques using machine learning. Based on the results of this analysis, we will look at how to extract the relative cultural index and present and analyze the experimental results derived by setting and training a model for predicting the cultural demand index. In this paper, we analyzed how to prepare objective guidelines and standards for implementing cultural supply policies based on big data.

## II. MATERIALS AND METHOD

### A. Concepts of Cultural Indicators

The concept of a cultural index emerged with the National Cultural Index System Development report published by the Ministry of Culture and Tourism in 2000, and the term cultural index appeared in the version of the same report published in 2002. Cultural indicators are numerical indicators that can explain culture [8]. In addition, a cultural index is a standardized comprehensive number defined for comparison, which allows various conditions related to culture to be compared to a reference point. Therefore, cultural indicators and indices can be used to compare regional cultural levels and serve as reference points [9].

To measure the cultural diversity index, cultural diversity index surveys and index measurement studies are conducted to collect various statistical data [10]. An index value that can represent cultural characteristics is then extracted from the data. This study investigated and classified detailed items for deriving cultural diversity index values that are difficult to capture using existing statistical data [11].

In addition, a cultural diversity index was extracted by considering the standardization of research results and weight values based on expert advice by conducting a survey on the status of cultural diversity and level diagnosis. In this manner, a comparative analysis study was conducted between countries and domestic cultural diversity domains and regions, and research results related to various indicators of cultural diversity were derived [12].

We attempted to derive support measures to solve social issues using big spatial data and establish effective policies for welfare, disaster, and real estate, which are important factors from a national policy perspective [12]–[14]. We examined the meaning of data-based policy establishment as a solution to using data in the policy process and analyzed domestic and foreign cases that have been or can be applied to the public policy process based on improved data processing and management technology [14]. Several other studies have been conducted on this topic. Studies have also shown how data can support and utilize rational public policy establishment [15]. Based on these efforts, big-data-related processing and analysis technologies have been researched. It has been determined that it is possible to provide information that can enable the value and standards of decision-making to

lead to highly effective policy decisions based on big data for policy establishment.

### B. Data-driven Policy Decision Making

Data-driven policymaking means making data-based decisions in the policy-making process [16]. The origin of the term ‘data-driven’ is ‘evidence-based’, which was introduced in the business, health, and medical field. This concept has emerged based on the importance of decision-making that faithfully and carefully uses objective information in making decisions [16].

This concept was introduced into the policy field by the “White Paper: Modernizing Government” published by the British government in March 1999, it is scientific rationality through securing a variety of objective grounds to check the individual and collective viewpoints [16]. In the United States, starting with the Obama administration, “The Obama evidence-based social policy initiative” was announced, and the “Commission on Evidence-Based Policymaking (CEP)” was established in 2017. In 2018, the “Foundations for Evidence-Based Policy-making Act” was enacted and the public sector is actively making efforts to create, manage, analyze, and revitalize policy-making grounds [16].

They are trying to collect and analyze data in order to implement a ‘data-driven’ policy rather than ‘simple evidence’ while defining the characteristics of available information using the scope of evidence [16]. In the South Korea, data is defined as ‘anything in a formatted or unstructured form that exists in a machine-readable form as information that is generated or processed through a device with information processing capability’, and based on the need for informatization and standardization As such, it is also being used in various government data utilization policies such as the financial business management system and administrative DB [16].

As a result, the meaning of the term ‘data-driven’ can be defined as ‘data’ with a purpose, created through systematic methods and statistical activities, and existing in a form that machines can process. These data types are classified slightly differently for each researcher, but examples include policy result information, descriptive surveys, administrative data, performance information, scenario and forecast data through surveys, economic feasibility-based information, and social justice and ethical evidence [16].

Policy decision-making is made at the policy evaluation and inspection stage, and feedback is provided through policy adjustment and modification [16]. Therefore, policy decision-making is a project tool to improve project efficiency and can take various forms, such as selection and change of support targets, budget increase and reduction to enhance policy effectiveness, the introduction of new projects, and conversion of the competent department [16].

### C. Prediction Techniques Using Machine Learning

Demand forecasting in governments and businesses is important in planning and managing production, materials, and logistics [17]. If demand is predicted well, waste of policy can be efficiently managed by product stockpiling in industries and warehouses, so it is possible to plan whether it is efficient to stockpile and estimate the logistics of moving products from production sites warehouses [17]. It is also

possible to plan for stockpiling from the purchase of raw materials needed to produce the product [17].

Since shipment planning for the manufactured products is also possible, demand forecasting is the most important part of a company's supply chain management (SCM) [17]. Demand forecasting is a field in which time series models that predict future demand are widely used by learning past demand trends [17]. As a traditional machine learning model in which the value to be predicted is affected by a past point in time, a moving average model that predicts the current value using an error or variation value that the model cannot explain, difference, branch regression, and moving average Autoregressive cumulative moving averages exist [17], [18].

These traditional time series models have one thing in common: they try to predict the future using only the patterns and errors of past data [17], [19]. Recently, in addition to historical data, various predictive models are being created that learn independent variables that affect demand at each historical point in time [17], [20]. A typical example is the ARIMAX model, which predicts future demand based on the independent variable at a future point in time by discovering the independent variable that affects the demand to be predicted [17], [21]. There is also LSTM, which is a kind of RNN using deep learning, and Prophet, Facebook's forecasting library, has a model developed into time series prediction in GAM [17], [22].

As such, various machine learning models for time series prediction exist, and it is important to identify, select, and apply the characteristics of the data along with the strengths and weaknesses of each model. The specific steps of the prediction technique using machine learning are as follows [17], [23], [24], [25].

1) *Data collection*: In the learning phase, it is necessary to collect a variety of learning data that is likely to have an impact on that point in the past, and the data must be able to be collected in the future.

2) *Create variable*: Using these variously collected learning data, the relationship between variables is identified through verification statistics and correlation coefficient. In addition, variables that can have a significant impact on demand are created by generating derived variables and understanding the relationship between the derived variables and the demand to be predicted [17], [23].

### 3) - *Learning*

Machine learning models such as time series models, representatively ARIMAX, Random Forest, eXtreme gradient boosting (XGBoost), and Light gradient boost machine (LGBM) regression, are trained with historical data in which various generated variables exist [17], [24]. Analysis methods using machine learning can utilize ensemble learning using multiple models rather than using a single model to increase predictive power [17], [25].

XGBoost, which is used in various fields due to its excellent performance, is also one of ensemble learning. XGBoost applied a boosting technique to create a stronger predictive model by combining the decision tree model with predictive models capable of simple classification [25]. The error is reduced by learning the given data through a classifier and learning the error appearing from the learned result in

another classifier. An algorithm that shows excellent performance when predicting structured data can be used in a regression model or a classification model [25].

It is faster than the GBM (gradient boosting machine) model, and various variables can be adjusted to prevent overfitting [25]. In general, it shows excellent predictive performance in classification and regression domains. The fact that XGBoost is useful for solving problems in various fields can disprove the usefulness of the model, XGBoost can quantify and interpret the importance of input variables through Shapley values for the interpretation of the developed model [25]. Shapley value is one of the algorithms of Game Theory, and it refers to the technique of calculating the contribution of each player in the game [25].

LGBM is an algorithm that increases accuracy by adding weights to misclassified data [25]. It is a tree-based algorithm, and it is a leaf-wise algorithm because it adds weight to only the wrong part. It has the advantage of high accuracy and less time to perform machine learning by performing leaf-central tree splitting (Leaf Wise) in the part with the maximum loss value [25].

4) *Evaluation*: Using the learned models, a specific period in the past is predicted as an evaluation target, and the model with the minimum error is selected as the optimal model by comparing the demand forecast value of each model with the actual demand value at that time [17], [24]. It is important to evaluate the prediction accuracy using the correct answer sheet, and it is not appropriate to use the size of the residual as an evaluation of prediction accuracy [17]. It is important to determine the accuracy of the predictions by how well the model fits the new data it was not used to training [17]. Various indicators use error to evaluate the accuracy of prediction. It is also possible to use a specific index among the evaluation indexes or use all indexes and the ranking summation for each index as the selection criterion for the optimal model [17]. It is an evaluation method that is not affected only by a specific period to evaluate the model using a different period as a validation set than the traditional evaluation method [17, 24].

5) *Prediction*: For each product, the model with the smallest error during the evaluation period is selected and future demand is predicted with the future data of the variables used for learning [17]. There is also a method to apply and predict the ensemble, unlike the method of evaluating the model using these various periods as a validation set [17]. In general, it is a method to find a model that predicts well only in a specific evaluation period; ensemble learning can be expected to reduce the prediction variance [17]. These ensemble techniques include Stacking, Voting, Bagging, and Boosting [17]. Typically, the Stacking ensemble is a technique for generating and learning separate training data for the final model to learn based on the prediction data generated by individual models trained with the training data [17], [24].

## D. *Related Research Trends*

A previous study on the establishment of a regional cultural policy measured relative cultural indices. In our study, only items related to cultural infrastructure were considered in terms of the supply of culture to obtain a relative cultural

index. The method for measuring relative culture indices used in previous studies is illustrated in fig. 1 [26].

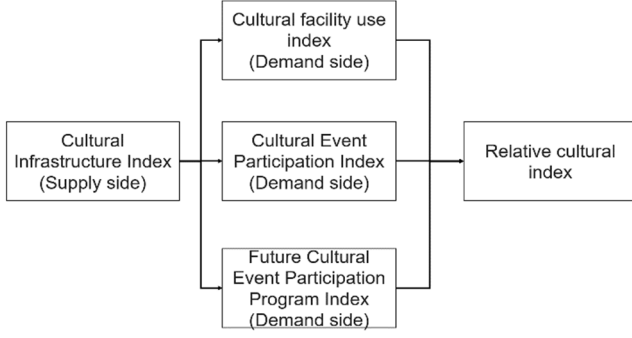


Fig. 1 Method used for measuring relative cultural indices in previous studies

However, this study determined that each region allocates a cultural budget differently. Therefore, the cultural budget items compared to the policy budgets of each region were also used to measure cultural supply indices. In addition, the Human Development Index (HDI) method proposed by the United Nations Development Program was adopted to standardize the level of comparison for calculating each index in previous studies [26, 27]. The formula for the HDI standardization method used in previous studies is written as follows:

$$HDI(i) = \frac{Actual(x_i) - Min(x_i)}{Max(x_i) - Min(x_i)} \quad (1)$$

However, this method tended to unify values to a similar level when constructing the predictive model used in our study. Therefore, our model was constructed without applying the HDI standardization technique. A log transformation was applied if the data distribution was abnormal at this stage. Therefore, in this study, in addition to comparisons using the calculated indices, we attempted to predict which policies are the most effective for revitalizing the cultural consumption of local residents. We propose our model as a helpful tool for effective policy establishment.

### III. RESULTS AND DISCUSSION

To make cultural policy decisions, existing regional cultural policy indicators, recent regional cultural resource indicators (cultural heritage, cultural infrastructure, cultural facility utilization, cultural resources), regional cultural activity indicators (activity support, activity status, activities, manpower), and local culture enjoyment indices (cultural enjoyment, cultural welfare) should be systematized to calculate a local culture index [28]. Among these factors, it is possible to categorize the target, criteria, and content of analysis and to analyze gaps in infrastructure based on cultural establishments such as museums, art galleries, and cultural centers, where public data can be accessed [29].

Recently, there have been many difficulties in engaging in real-world culture as a result of the impact of COVID-19, so by collecting and analyzing data from the pre-COVID era, participation in culture and enjoyment of culture can be improved in the post-COVID era.

Therefore, this study calculated a culture supply index and a culture demand index to construct a model supporting cultural supply policy decision-making. Subsequently, a

relative cultural index was calculated using the previously calculated supply and demand indices, and the cultural level of each region was compared. We propose selecting underdeveloped regions based on relative cultural indices and prioritizing policy implementation in these regions.

In addition, this study assumed that the success of a culture supply policy improves the cultural demand index. Therefore, we propose selecting cultural supply items that have the greatest influence on the demand index for each region and establishing supply policies for those items. Furthermore, by using the proposed model, we quantify the extent to which the culture demand index will be improved or degraded when a culture supply policy is implemented. Accordingly, our method can be used as an objective decision-making tool for implementing policies.

#### A. Relative Culture Index

A relative cultural index was calculated by considering the cultural supply-side index (cultural infrastructure index) and cultural demand-side index (cultural facility use index, cultural event participation index, future intention of culture). The units of supply for calculating the relative cultural index are all different, and therefore, standardization must be performed to calculate each index.

First, we used the cultural infrastructure index to calculate the supply-side index. The number of cultural facilities in each region was divided by the local population and used to calculate the supply index. We aimed to determine the number of cultural facilities allocated per person in each region based on these calculations. Second, the ratio of the cultural budget to the total budget in each region was calculated. Before adding the relevant items, the correlation between the cultural demand indices of each region calculated using the method presented in this paper and the culture-related budget ratio to the total budget was analyzed.

As a result, we confirmed a positive correlation in which the cultural demand index increased when the value of the relevant items increased. Therefore, we can use these items to calculate relative culture indices. The cultural facility use index, cultural event participation index, and future culture intention index are used to calculate the demand-side index. This study used the cultural demand index as a cultural supply policy establishment index. Based on the model presented in this paper, we propose a policy of supplying items that can significantly increase the cultural demand index. However, in general, it is suggested to reduce the cultural demand index and refrain from policies related to supply items that do not change it significantly.

#### B. Model Training

In this study, the cultural demand index was used as an index for policy establishment based on the efficient supply of culture. Therefore, we constructed a model to predict the cultural demand index when implementing specific cultural supply items policies. To predict the cultural demand index according to the cultural supply, we use the cultural demand index data based on data related to cultural supply from the last five years prior to COVID-19 in each region in Gyeonggi-do.

We utilized a public data portal and the Culture Data Plaza operated by the Ministry of Culture, Sports, and Tourism in

Korea for data collection. Datasets such as surveys and culture-related budget status were collected, and because there was significant variation in the data on cultural facilities across the country, data were collected to calculate relative cultural indices by traveling. Items related to the supply index used in this study included the number of cultural infrastructures by region and the ratio of the cultural budget to the total budget. However, these data exhibited a skewed distribution. Therefore, we applied a log transformation to adjust the distribution.

Next, there were three items related to the demand index: use of cultural facilities, participation in cultural events, and intention to enjoy culture in the future. However, among these three items, the intention to enjoy culture in the future was a categorical variable. Therefore, the mean encoding method, which performs statistical calculations with target values based on categorical variables, was applied. Because the items related to the demand index also exhibited a somewhat skewed distribution, log transformation was applied to adjust the distribution.

This study applied two algorithms, namely XGBoost and the light gradient boosting machine (LGBM). A final policy establishment decision support model was constructed by applying an ensemble method using these two models. Subsequently, a five-fold cross-validation technique was adopted as a validation technique to evaluate model performance. The mean absolute percentage error and accuracy were used as evaluation metrics.

### C. Experimental Results

The cultural index was calculated for each region in terms of supply, as shown in Table 1, Seoul and Gyeonggi-do were distributed in the upper ranks. Also, when looking at the demand index, as shown in Table 2, Seoul and Gyeonggi-do are similarly distributed at the top. The relative cultural index is expressed by synthesizing the above two supply and demand indices shown in Table 3. Seoul and Gyeonggi-do are at the top, and Daejeon and Chungcheong-do are at the bottom. By using the model trained in this study, the city of Daejeon was selected as a sub-region to predict which cultural supply items could significantly improve the cultural demand index. The regression coefficient of the theater item was the greatest at 0.4375 among the items related to supply.

TABLE I  
SUPPLY INDEX TOP DISTRIBUTION CITIES

City	Supply
Seoul	0.8958
Gyeonggi-do & Incheon	0.8646
Busan & Gyeongsang-do	0.7917
Gwangju & Jeolla-do	0.5938
Daejeon & Chungcheong-do	0.4063
Gangwon-do & Jeju-do	0.3542

TABLE II  
DEMAND INDEX TOP DISTRIBUTION CITIES

City	Demand
Seoul	0.5628
Gyeonggi-do & Incheon	0.5246
Gwangju & Jeolla-do	0.5017
Daejeon & Chungcheong-do	0.4642
Gangwon-do & Jeju-do	0.4888
Busan & Gyeongsang-do	0.4765

TABLE III  
RELATIVE CULTURAL INDEX COMPARISON RESULTS

City	Demand
Seoul	0.6844
Gyeonggi-do & Incheon	0.5997
Gwangju & Jeolla-do	0.5428
Gangwon-do & Jeju-do	0.5084
Busan & Gyeongsang-do	0.4739
Daejeon & Chungcheong-do	0.4375

In the final report of the 2020 Living Culture Survey and Effectiveness Study on the COVID-19 era, the main active life culture fields were literature (13.5%) and outdoor activities (12.6%), whereas living culture awareness was a participatory and subjective cultural activity (31.4%). This result can be viewed in the same context as those for voluntary cultural activities (29.1%), viewing/listening cultural activities (24.6%), varying cultural activities (9.9%), and community-connected activities (4.4%), and others (0.5%).

If people can discover and quantify the life culture that they most hope for in their current scenario and predict its availability, it will be possible to fulfil people's desires and improve their value of daily life, sense of happiness, and quality of life. When the model for predicting the cultural demand index designed in this study was evaluated, it did not achieve good performance, and this is because data from the five years before COVID-19 were not sufficient for training the model. However, if additional data can be obtained and utilized in the future, performance improvements will be possible.

### IV. CONCLUSION

In the Republic of Korea, a data-related law was passed at the plenary session of the National Assembly in January 2020. As a result, the problem of personal information protection, which has been a limitation in collecting information based on beneficiary standards and analyzing causal effectiveness through the integration of administrative DBs in the public sector, has been found to be solved by using pseudonymous information. Now, statistical analysis and utilization for the public interest are institutionally possible by integrating data that has safely processed personal identification information.

If data analysis for data-based policy decision-making is performed through such institutional use, effect-based policy performance analysis becomes possible in the existing indicator-oriented performance management system. As a result, the link between performance information and budget information is visualized, and it is expected that significant progress will be made in policy decision-making, such as policy adjustment and budget investment prioritization.

Recently, several studies have been conducted to solve social problems based on data. Following this approach, this paper proposed a model for effective cultural policy establishment based on cultural data and demonstrated that rational decision-making support is possible based on such a model. In this study, the human development index was calculated using machine learning, and the relative culture index was calculated based on the calculated results.

When establishing future cultural policies by calculating the relative index based on the result of calculating the relative cultural index. It is conducted by introducing the prediction

technique of machine learning based on existing data and the supply index and demand index for each city in the COVID-19 situation. As a result, it was analyzed so that it could be used as a basis for establishing future cultural policies.

By using the proposed method, it is expected that more diverse decision-making can be rationally supported if the analysis is conducted nationwide rather than only in Gyeonggi-do. Furthermore, when selecting predictive variables, better performance is expected if variables with high explanatory power are discovered and added through the process of considering variables related to other cultures.

## REFERENCES

- [1] H.-Y. Choi, E.-H. Seo, and H.-J. Jun, "Social Diversity and Quality of Life among People Living in Seoul: An Application of Multilevel Model," *Journal of the Korean Regional Science Association*, vol. 36, no. 1, pp. 69–88, Mar. 2020. DOI : 10.22669/kras.2020.36.1.069.
- [2] Ahn, P., "Role of Senior Program for Financial Independence in Culture and Art Facility," *Journal of Korean Society for Quality Management*, vol. 48, no. 2, pp. 345-359. 2020. DOI : 10.7469/JKSQM.2020.48.2.345.
- [3] Jeon, C. H., Kim, D. J., Kim, S. K., Kim, D. J., Lee, H. M., & Park, H. J., "Validation in the cross-cultural adaptation of the Korean version of the Oswestry Disability Index," *Journal of Korean medical science*, vol. 21, no. 6, pp. 1092-1097, 2006. DOI: 10.3346/jkms.2006.21.6.1092.
- [4] Kim, B. R., Shin, J., Guevarra, R., Lee, J. H., Kim, D. W., Seol, K. H., & Isaacson, R. E., "Deciphering diversity indices for a better understanding of microbial communities," *J Microbiol Biotechnol*, vol. 27, no. 12, pp. 2089-2093, 2017. DOI: 10.4014/jmb.1709.09027.
- [5] Kim, H. Y., & Cho, J. S., "Data governance framework for big data implementation with a case of Korea," In 2017 IEEE International Congress on Big Data (BigData Congress), pp. 384-391, 2017. DOI: 10.1109/BigDataCongress.2017.56.
- [6] Engin, Z., & Treleaven, P., "Algorithmic government: Automating public services and supporting civil servants in using data science technologies," *The Computer Journal*, vol. 62, no. 3, pp. 448-460, 2019. DOI: 10.1093/comjnl/bxy082.
- [7] Chen, J., Chen, Y., Du, X., Li, C., Lu, J., Zhao, S., & Zhou, X., "Big data challenge: a data management perspective," *Frontiers of computer Science*, vol. 7, no. 2, pp. 157-164, 2013. <https://doi.org/10.1007/s11704-013-3903-7>.
- [8] Chun, S., Shulman, S., Sandoval, R., & Hovy, E., "Government 2.0: Making connections between citizens, data and government," *Information Polity*, vol. 15, no. 1-2, pp. 1-9, 2010. DOI 10.3233/IP-2010-0205.
- [9] Vanderlinde, R., Dexter, S., & van Braak, J., "School-based ICT policy plans in primary education: Elements, typologies and underlying processes," *British Journal of Educational Technology*, vol. 43, no. 3, pp. 505-519, 2012. <https://doi.org/10.1111/j.1467-8535.2011.01191.x>.
- [10] Visvizi, Anna, et al., "Policy making for smart cities: Innovation and social inclusive economic growth for sustainability," *Journal of Science and Technology Policy Management*. 2018. <https://doi.org/10.1108/JSTPM-07-2018-079>.
- [11] Park, E., "Positive or negative? Public perceptions of nuclear energy in South Korea: Evidence from Big Data," *Nuclear Engineering and Technology*, vol. 51, no. 2, pp. 626-630, 2019. <https://doi.org/10.1016/j.net.2018.10.025>.
- [12] Kim, E. S., Choi, Y., & Byun, J., "Big Data Analytics in Government: Improving Decision Making for R&D Investment in Korean SMEs," *Sustainability*, vol. 12, no. 1, p. 202, 2020. <https://doi.org/10.3390/su12010202>.
- [13] Olszewski, R., Pałka, P., Wendland, A., & Majdzińska, K., "Application of cooperative game theory in a spatial context: An example of the application of the community-led local development instrument for the decision support system of biogas plants construction," *Land Use Policy*, 108, 105485, 2021. <https://doi.org/10.1016/j.landusepol.2021.105485>.
- [14] Ordóñez de Pablos, P., & Lytras, M., "Knowledge management, innovation and big data: Implications for sustainability, policy making and competitiveness," *Sustainability*, vol. 10, no. 6, 2018. DOI : 10.3390/su10062073.
- [15] Yoon, D., "The information science policy for the public open data of the national research institute," *Cogent Business & Management*, vol. 4, no. 1, 1406321, 2017. DOI : 10.1080/23311975.2017.1406321.
- [16] S. Kim, "A Study on Development of Policy Attributes Taxonomy for Data-based Decision Making," *The Journal of Information Systems*, vol. 29, no. 3, pp. 1–34, Sep. 2020. DOI : 10.5859/KAIS.2020.29.3.1.
- [17] LG CNS, Difficulty identifying demand, can it be predicted with machine learning?, <https://post.naver.com/viewer/postView.nhn?volumeNo=29283830&memberNo=3185448&vType=VERTICAL>, (visited Nov. 2021).
- [18] R. B. Caldo and J. Palomaria, "Automated Driver's Assessment and Driving Simulator System," *International Journal of Advanced Science and Convergence*, vol. 1, no. 1, pp. 23-31, 2019. DOI: 10.22662/IJASC.2019.1.1.023.
- [19] J. Lee and C. H. Lee, "A Study on Changes in Foreign Tourists Using Big Data Analysis Method," *International Journal of Advanced Science and Convergence*, vol. 2, no. 4, pp. 25-30, 2020. DOI: 10.22662/IJASC.2020.2.4.025.
- [20] Nishant, Rohit; Kennedy, Mike; Corbett, Jacqueline, "Artificial intelligence for sustainability: Challenges, opportunities, and a research agenda," *International Journal of Information Management*, 53: 102104, 2020. <https://doi.org/10.1016/j.ijinfomgt.2020.102104>.
- [21] Modgil, Sachin; Gupta, Shivam; Bhushan, Bharat, "Building a living economy through modern information decision support systems and UN sustainable development goals," *Production Planning & Control*, 31.11-12, pp. 967-987, 2020. <https://doi.org/10.1080/09537287.2019.1695916>.
- [22] Constantiou, Ioanna; Shollo, Arisa; Vendelo, Morten Thanning, "Mobilizing intuitive judgement during organizational decision making: When business intelligence is not the only thing that matters," *Decision Support Systems*, 121: pp. 51-61, 2019. <https://doi.org/10.1016/j.dss.2019.04.004>.
- [23] Kim, Yeonsoo; OH, Jooseok; Kim, Seiyong, "The Transition from Traditional Infrastructure to Living SOC and Its Effectiveness for Community Sustainability: The Case of South Korea," *Sustainability*, 12.24: 10227, 2020. <https://doi.org/10.3390/su122410227>.
- [24] Dahlgaard-Park, Su Mi; Reyes, Lidia; Chen, Chi-Kuang, "The evolution and convergence of total quality management and management theories," *Total Quality Management & Business Excellence*, 29.9-10, pp. 1108-1128, 2018. <https://doi.org/10.1080/14783363.2018.1486556>.
- [25] Abou Omar, Kamil Belkhat, "XGBoost and LGBM for Porto Seguro's Kaggle challenge: A comparison," *Preprint Semester Project*, 2018.
- [26] Castelnovo, Paolo, Valentina Morretta, and Michela Vecchi, "Regional disparities and industrial structure: territorial capital and productivity in Italian firms," *Regional Studies*, vol. 54 no. 12, pp. 1709-1723, 2020. <https://doi.org/10.1080/00343404.2020.1763941>.
- [27] Kim K., Hong S., "Analysis on Regional Disparities in Culture & Arts Infrastructure: Focused on Gyeonggi Province," *The Journal of Korean Policy Studies*, vol. 21, no. 2, pp.27-50, 2021.
- [28] Culture Data Plaza", Culture.go.kr, 2021. [Online]. Available: <https://www.culture.go.kr/>. [Accessed: 03- Nov- 2021].
- [29] Public data portal ", data.go.kr, 2021. [Online]. Available: <https://www.data.go.kr/>. [Accessed: 03- Nov- 2021].