

Interactive Recommendation AI to Support Transparent Human Decision Making

Yuri Nakao Kotaro Ohori Hirokazu Anai

AI technology which has been increasingly used in recent years has a disadvantage; it is difficult to explain the reasons behind its outputs in understandable ways to humans. To address this problem, research on explainable AI (XAI) is underway. However, it is necessary to enable XAI to interact with its users to support human decision making. As a prototype of interactive XAI, Fujitsu Laboratories has developed AI to recommend relocation destinations with the cooperation of the local government of Itoshima City, Fukuoka Prefecture in Japan. This is a recommender system that supports users in urban areas with decision-making when selecting suitable candidate relocation destinations. In response to the preferences input by users, the system provides not only lists of the candidate relocation destinations but also information for the users to learn about the destinations in detail. This article describes an interactive AI that realizes collaboration with humans. It also presents the results of a subject experiment on those who wish to relocate to Itoshima City and the future issues and direction.

1. Introduction

It has been a long time since AI entered its third boom. At present, mainstream AI such as deep learning takes data-driven approaches, meaning that AI analyzes collected data for the purpose of prediction and decision-making. This kind of AI has both advantages and disadvantages. As an advantage, it can extract feature values required to improve the accuracy of the outputs that humans cannot mention clearly. On the other hand, this kind of AI has a disadvantage that its outputs and the extracted feature values sometimes cannot be interpreted with human intuition. In the context of business and daily lives, the reasons must be explained in an understandable manner for humans, as do the outputs from the AI themselves. However, it was difficult for conventional AI to provide such explanations.

In order to address this issue, explainable AI (XAI) technology has been developed [1]. XAI is intended to output the reasons for the results from AI in a way that humans can understand. For example, one famous technology is developed to visualize which part of an image was effective as the reason for the judgment by AI [2].

To support decision making in the context of business and daily lives, however, developing only XAI is not sufficient. It is necessary to develop a technology that enables users to evaluate the outputs from AI critically and to modify the results. This causes users to think over seriously in the course of interaction before taking actual action. Only when the process of critical interaction is confirmed can AI collaborate with humans. **Figure 1** shows the difference between the conventional approach and this approach. Conventionally, XAI has been developed focusing on how effective AI will persuade humans with the displayed explanations. On the other hand, our approach focuses on how to ensure a situation where users make decisions after due consideration through thinking about the outputs from AI and giving critical feedback to the outputs.

In order to ensure this situation, the AI technology itself must be built in a way that can be understood by the users. Developing interfaces and the technology itself in a way that lets users easily understand the mechanism of the AI is referred to as being “transparent.” Development of interactive XAI that allows humans to make decisions while understanding the mechanism of AI through trial

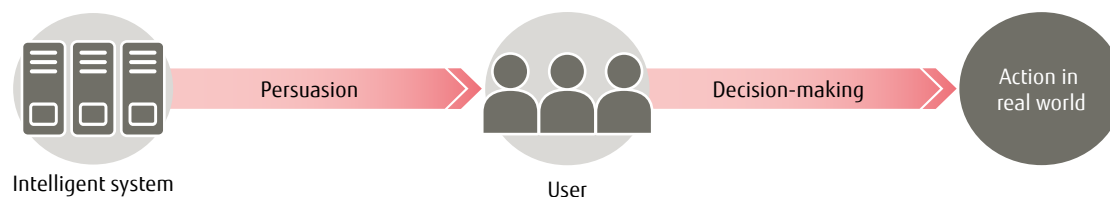
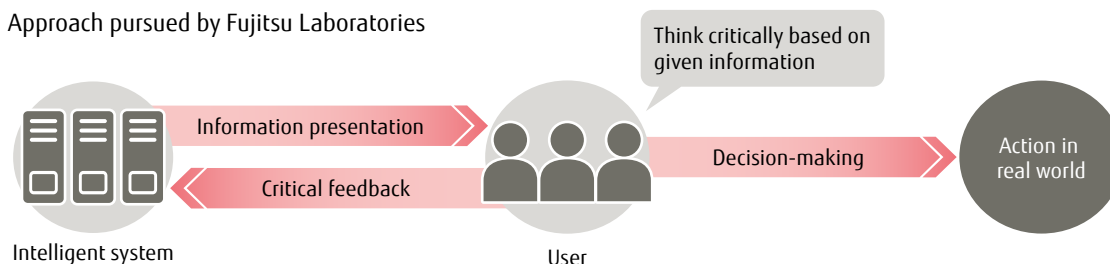
Conventional AI approach**Approach pursued by Fujitsu Laboratories**

Figure 1
Differences between conventional approaches and this approach.

and error leads to the realization of transparent and trustworthy AI.

This article presents a case of recommendation AI for relocation destinations, which was jointly developed and field-tested by Itoshima City, Fukuoka Prefecture in Japan and Fujitsu Laboratories. This AI was developed to investigate the effects of explanation given by AI on users and their reactions to the explanation. First, this article explains the conventional technologies and their problems. Next, by taking the recommendation AI for relocation destinations as an example, how the approach mentioned above is implemented as an actual service is shown. Finally, the contents of the system evaluation experiment and future direction are described.

2. Conventional technologies and their problems

There have been systems to support human decision-making through interactions with humans. Academically, in the latter half of the 2000s, a technology called 'critiquing' was developed [3]. The technology was built by adding a function of feedback from users to recommender systems for products such as automobiles and digital cameras [3]. In a business context, recommender systems with which users can filter items based on user-set conditions are also

widespread among services such as job matching and real estate matching [4]. However, while these systems can receive feedback from users, they do not focus on making the technology itself transparent. The more important a decision to be made is, the more it is necessary for the user to thoroughly understand why the AI is providing the recommendations. To this end, it is essential to improve the transparency of the technology itself. For the system presented in this article, we develop an interactive AI whose mechanism is as transparent to the users as possible.

3. Relocation destination recommendation AI to support human decision-making

This section presents an interactive recommender system called a relocation destination recommendation AI. This system is developed to provide a service that assists people living in urban areas in Japan who want to relocate to more rural provinces with finding areas that suit themselves. The system is developed through joint research by the local government of Itoshima City, Fukuoka Prefecture in Japan and Fujitsu Laboratories. Urban residents lack prior knowledge of the detailed information concerning rural areas. Therefore, it is necessary for them to learn about the potential relocation destinations while searching for an area to which they

want to relocate. This system recommends suitable areas in Itoshima City to live in for those who want to relocate.

First, Fujitsu Laboratories conducted multiple interviews with people who had already moved from urban areas to Itoshima City and with those who were considering moving there in the future. The results of interviews were used to identify the features of the areas that urban residents care about when deciding on a destination in which to relocate. Next, those features were quantified based on regional data. With the data, regional indicators corresponding to the features of the areas such as traffic, shopping, and hospitals are created. At the same time, factors to classify people who want to relocate to Itoshima City were identified based on the results of the interviews. The factors include whether people can drive cars and the age brackets of their children. People who want to relocate are classified into some tens of categories. In addition,

an online study to ask quantitative questionnaires is conducted to define the degrees of importance (i.e., to what extent people consider the regional indicator important) for each regional indicator from each category of people. Through this process, we have developed a system equipped with the GUI shown in **Figure 2**.

In this GUI, the category of the user who wants to relocate is shown in the upper left, and the regional indicators regarded as important by people in the same category as the user are shown in the upper right. The degrees of importance for the regional indicators for the user are shown by the bar graphs under Step 1 in Figure 2. By clicking on the arrows above and below the bar graphs, this GUI enables users to change the degrees of importance for regional indicators. Through this function to modify the degree of importance, the users can specify the appropriate balance of degrees of importance as they wish. Under the bar graphs, the ranking of recommended relocation destinations—not

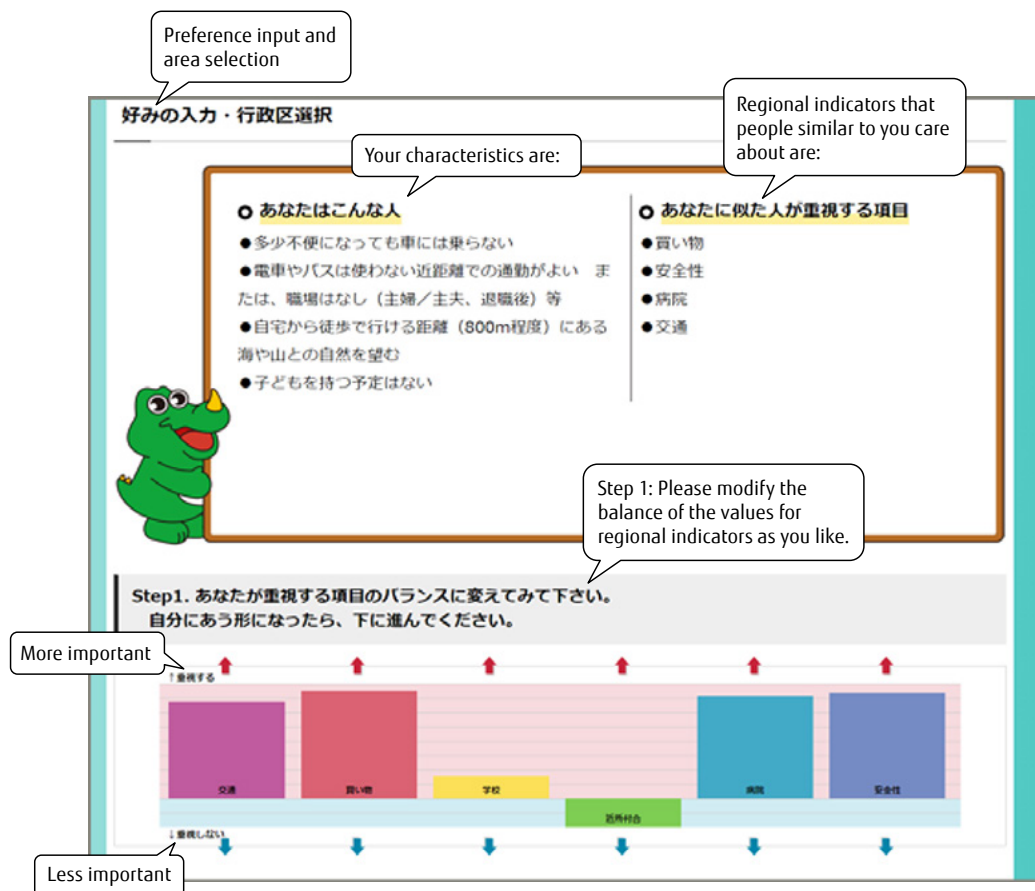


Figure 2
GUI for the Itoshima City relocation recommendation AI system.

shown in Figure 2—is shown together with the values of regional indicators that the areas have. Users can adjust the degree of importance with the bar graphs to show their preferences while learning the information.

Additionally, this system has a mechanism to learn interactions between the system itself and its users. Based on the results of the learning, the system automatically modifies the user category. This allows the system to learn users' characteristics as more users use it, as well allowing users to learn information about the recommended areas.

To ensure transparency, this system uses a linear model based on multi-attribute utility theory [5], which is modeled initial phase of human decision-making. In other words, the score of the individual area is the sum of the products of the regional indicator value given to each area and the degree of importance for the indicator set by the user, which are shown by the bar graphs.

4. System evaluation

The effect of this system was investigated using a subject experiment. The subject experiment was conducted on a total of 99 participants, including 16 people visiting in information sessions to introduce Itoshima City to those who are considering relocations to rural areas and 83 participants recruited on the Internet who do not have prior knowledge of Itoshima City but who wish to relocate to rural areas. Of the participants recruited on the Internet, 55 underwent the experiment with a system equipped with the initial degree of importance created based on the quantitative questionnaires. The remaining 28 participants underwent the experiment with a system that has trained with the data from the 55 participants in the previous phase.

To evaluate the effect of the training of the system about the users' behavior, we compared the system log data of the participants before training to that after training. The analyzed data compared includes the distance that participants' preferences traveled in the preference space and the number of times the participants renewed the list of recommended areas by changing their preferences. Their preferences mean the degrees of importance for regional indicators shown by the lengths of the bar graphs in Figure 2. A preference space refers to a space where the degrees of importance given by the user to the respective regional

indicators are plotted. Because we define six regional indicators, the preference space is composed as a six-dimensional space.

As the results, we found that participants who visited the information session on Itoshima City tended to be reluctant to change their preferences. One participant commented that he felt that the system told him to "live here."

Participants recruited on the Internet experience the process of the consultation about the relocation to Itoshima City with the interactive AI system. First, participants receive recommendations of the relocation destinations from our system. Next, they discuss where to relocate with staff from the local government of Itoshima City.

Figure 3 shows quantitative results from the experiment with the participants recruited on the Internet. Figure 3 (a) shows the results of a comparison of the average value of the travel distances in the preference space between before and after training. While the average before training is 101.1, the value after training is 74.2 and the p-value [6] is 0.003. This indicates that the travel distance decreases significantly after training. Figure 3 (b) shows the result of the number of renewals of the recommendation list. The average value before training is 3.8, whereas the average value after training is 3.5 and the p-value was 0.55. This shows that there is no significant difference between before and after training. While the number of the participants is very different: 55 before training

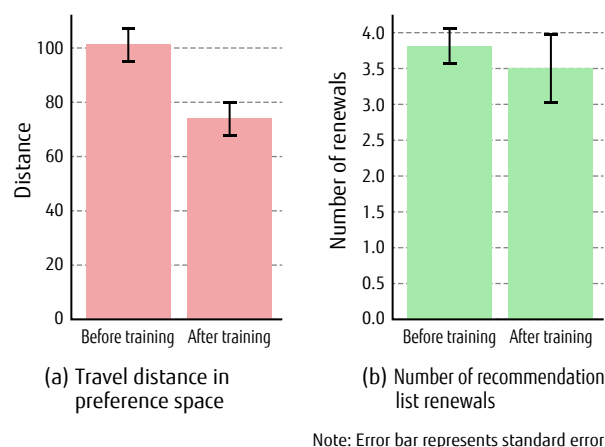


Figure 3
Travel distance in preference space and number of recommendation list renewals.

and 28 after training, we found a tendency that the number of renewals of the lists of the recommendation by the participants did not decrease even if the travel distances in the preference space decreased. In other words, the participants attempted to change their preferences approximately four times on average even if the areas that meet their requirements were shown immediately. This indicates the tendency that people wanted to renew the recommendation list a certain number of times to find an area that suits them.

While using the system and in the questionnaires after the experiment, related to the advantages of the AI system for relocation consultations, participants commented that "It is easier and faster than consulting people (to obtain the answer from the AI system)," "(With the system,) I think I can find areas that are difficult to notice," and "(The system) provides impartial and objective information." Meanwhile, disadvantages of the system are mentioned, such as "There are limits to text information when I want to know what living (in Itoshima City) is really like."

The participants also received information from staff in charge of relocation from the local government of Itoshima City. For the consultation with the staff, only positive points were mentioned such that "They offered many specific proposals from various angles related to living in Itoshima City" and "They answered immediately, even when asked complicated questions"; no negative points were mentioned.

This system was put to practical use on the website of the local government of Itoshima City in the form of the Itoshima City Relocation AI Recommendation System from September 2017 to April 2018 after the experiment. This resulted in the relocation of multiple households.

5. Future issues and direction

One future issue is the development of an interface that encourages users to interact with AI systems. Although we tried building a system to allow users to make decisions after interacting with the AI, the participants in the information sessions tended to regard the recommendations given by the AI as fixed instructions and were not willing to interact with the AI. This implicates that it is necessary to improve the interface to encourage more interaction.

In addition, although we used a linear model to

improve transparency, users found it difficult to grasp even the simple linear model [7]. This poses the issue of how to explain the model itself in an easy to understand way.

The potential of design orientation of development toward the interactive XAI described in Section 1 to achieve AI that can support appropriate human decision-making is not limited to the relocation destination recommendation AI presented in this article. The same approach can be used to design recommendation AI applicable to various fields such as tourism, real estate, and education.

6. Conclusion

This article described the necessity of interactive XAI to support decision-making in business and daily lives and the system to recommend relocation destinations developed as a prototype jointly with Itoshima City, Fukuoka Prefecture in Japan together with the results of the subject experiment. Fujitsu Laboratories intends to continue to work on the R&D of interactive XAI systems capable of supporting appropriate human decision-making through proper design of contact points between humans and technology, thereby realizing trustworthy AI.

All company and product names mentioned herein are trademarks or registered trademarks of their respective owners.

References and Notes

- [1] A. Abdul et al.: Trends and Trajectories for Explainable, Accountable and Intelligible Systems: An HCI Research Agenda. CHI '18 Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems, Paper No. 582 (April 21–26, 2018).
- [2] M. T. Ribeiro et al.: "Why Should I Trust You?: Explaining the Predictions of Any Classifier." KDD '16 Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp. 1135–1144 (August 13–17, 2016).
- [3] L. Chen et al.: "Critiquing-based recommenders: survey and emerging trends." User Model User-Adap Inter, Volume 22, Issue 1–2, pp. 125–150 (April 2012).
- [4] C. He et al.: "Interactive recommender systems: A survey of the state of the art and future research challenges and opportunities." Expert Systems with Applications, Volume 56, pp. 9–27 (September 2016).
- [5] W. Edwards et al.: Decision Technology. Annual Review of Psychology, Volume 52, pp. 581–606 (February 2001).

- [6] In statistical hypothesis testing, the p-value is the probability that a test statistic will be the value under a given hypothesis. In this article, we hypothesize that the average values of the measured quantities obtained from two unpaired groups are the same and use the value to indicate how unlikely it is to occur. The smaller the p-value, the less likely it is that the averages of the two groups are the same. In this article, a p-value smaller than 0.05 is regarded as a significant difference between the averages of the two groups.
- [7] Y. Nakao et al.: Generation of Hints to Overcome Difficulty in Operating Interactive Recommender Systems. IntRS '19: Joint Workshop on Interfaces and Human Decision Making for Recommender Systems (September 19, 2019).



Yuri Nakao

Fujitsu Laboratories Ltd., Artificial Intelligence Laboratory
Mr. Nakao is currently engaged in research on recommender systems, fairness-aware machine learning, and user observation for accessibility.



Kotaro Ohori

Fujitsu Laboratories Ltd., Artificial Intelligence Laboratory
Mr. Ohori is currently engaged in research on social system design using artificial intelligence and mathematical technology.



Hirokazu Anai

Fujitsu Laboratories Ltd., Artificial Intelligence Laboratory
Mr. Anai is currently engaged in research and development of formula manipulation, optimization, and theory and application of artificial intelligence.

This article first appeared in Fujitsu Technical Review, one of Fujitsu's technical information media. Please check out the other articles.

Fujitsu Technical Review

<https://www.fujitsu.com/global/technicalreview/>

