

Well-being Data Origination Using MROCs with Variable Quest: A Case Analysis of Gloom during COVID-19 Pandemic

Teruaki Hayashi*¹, Yumiko Nagoh¹, Kai Ishikawa²,
Hirohiko Ito², Kenichiro Tsuda², Yukio Ohsawa¹

¹The University of Tokyo, ²NEC Corporation
*hayashi@sys.t.u-tokyo.ac.jp

Abstract

Technologies using artificial intelligence (AI) have been implemented as services to solve various social problems. However, the contributions of AI to people's mentality and unknown/unobserved events have not been extensively discussed. In this study, we focus on people's mental changes caused by the COVID-19 pandemic and discuss the origin of data sources for well-being using marketing research online communities (MROCs) and variable quest (VQ). In the experiment, we selected 15 females aged between 20 and 40 who were interested in exploring how daily life has changed since the emergence of COVID-19 using MROCs. The analysis results by VQ revealed that the variable sets of the events differed with the situations, mental states, and attitudes, while not being featured in any of the MROC topics as keywords. The result suggests that abstracting the features of unobserved events as variable sets, can help us acquire information potentially contributing to unexplored data discovery for human well-being from texts not containing any information related to the data.

Introduction

AI technology can provide effective and automated solutions for issues that manifest in the physical world. However, AI's contribution to people's mentality and unknown/unobserved events has not been sufficiently investigated. Various industries worldwide have been severely affected by the coronavirus since the end of 2019, revealing gaps between social systems and enforcing major transformations in our lives. Data consisting of accumulated records of past events may not be effective in calculating infection risk or preventing wider infections. Moreover, not much is understood regarding the emotional issues that accompany changes in daily life.

This research considers accomplishing well-being data origination of unexplored data. Data origination is the act of data design/acquisition/utilization that reflects the subjective knowledge and diversity of perspectives of humans and aims to elucidate as well as support this process (Hayashi, 2020). Unexplored data signify a source of data containing

requests, imaginary parts, and events yet to be converted into data, or, the potential data of unobserved events. In this study, we used marketing research online communities (MROCs) and variable quest (VQ) to identify unexplored data on the gloom brought on by the COVID-19 pandemic, while discussing the origin of data sources for well-being.

Methods

MROCs are research methods for extracting consumer insights by building a closed online community of people who are knowledgeable about or interested in specific themes (Baldus, 2015). Unlike an ordinary questionnaire survey, this method obtains more substantial opinions by using the communication between participants to shed light on the relationships between their issues and the themes. VQ is a system for estimating variables that serve as compositional elements of data (Hayashi, 2017). Variables are data attribute sets, and one of the main compositional elements of the data. The advantage of using VQ in this study is that the sets of variables relevant to the event in question can be accessed from data-related collective intelligence by providing a text outline of an event that has not yet been noticed, or observed, even when keywords related to variables are not included.

We selected 15 females aged between 20 and 40, who were interested in exploring changes in daily life since the emergence of COVID-19 and received 314 comments regarding seven topics with MROCs. To understand how families, friends, and workplaces have changed, we chose the comments of the three topics as follows:

- Topic 2: Feelings about lifestyle changes due to the COVID-19 pandemic (66 comments)
- Topic 5: How cohabiting with relatives, colleagues, and friends has changed (30 comments)
- Topic 6: How feelings have changed after spending half a year living with COVID-19 (25 comments)

Table 1: Example comments and variable sets for each topic

#	Example comments	Example variable sets
Topic 2	<ul style="list-style-type: none"> While most companies allowed working from home, it was scary to go into office, and I felt different from everyone else. [...] I was haunted by this constant anxiety, so even on my days off, I couldn't get much emotional respite. I avoided behaviors that seemed to harm my physical condition [...] and became sensitive to changes in my body. 	"Attitudes about health," "work-related attitude," "attitudes about money," "occupation," "number of steps," "age," "number of cases," "number of deaths," "number of suicides," "prefecture," ...
Topic 5	<ul style="list-style-type: none"> Even if you're asking about something small, putting it into writing in an email or chat takes a surprisingly long time. More contact via email and phone, and more pressure to respond from family, friends, colleagues—everyone in general. Maybe it's feelings of unease, going out less, and spending more time at home. 	"Annual salary," "sales," "labor hours," "affiliation," "email address," "email sender," "occupation," "number of steps," "age," "number of cases," "number of deaths," "number of suicides," "prefecture," ...
Topic 6	<ul style="list-style-type: none"> Even in my daily life, I'm thinking about the future, a career change, studying; it feels like everything is looking toward the future. I used to choose where to go based loosely on hyped-up places and popular events, but since we've begun living with the coronavirus, whenever I need to run an errand, I actively go to places with fewer people. 	"Occupation type," "author name," "knowing about an event in advance," "purchase history," "birthday," "attitudes about health," "work-related attitude," "attitudes about money," ...

Results and Discussion

Topic 2 revealed many negative lifestyle changes during the COVID-19 pandemic; in particular, many participants referred to changes in their physical condition, daily habits, and the shift to remote working (Table 1). This suggests that it is important to capture these variables when observing how work habits have changed, including in remote working. Moreover, a variable group was considered important for grasping changes in attitudes toward work and lifestyles. For example, "attitudes about health," "work-related attitudes," and "attitudes about money" were found to be a particularly important variable for grasping and understanding the state of mental health. Furthermore, "occupation," "number of steps," and "age" can provide an idea of the context of insufficient exercise as a result of remote working. Next, "number of cases," "number of deaths," "number of suicides," and "prefecture" form a group of variables that reflect news about the alleged rise in suicide and daily infection rates during the pandemic. These variables provide more quantifiable data than the other variables.

Topic 5 deals with variables about changes in participants' personal circles and shared the same variables as remote working and news in Topic 2. Compared with Topic 2, the variables for Topic 5 did not concern psychological issues, instead centering on events occurring in the physical world, such as "annual salary," "sales," and "labor hours".

Topic 6 discussed the changes since the first declaration of emergency in Japan in April 2020. There were still many comments regarding the discomfort regarding work and daily life; however, contrary to Topic 2, there were many positive statements about accepting societal changes, beginning new activities, and gradually adjusting feelings. Accordingly, such variables appeared as "occupation type," owing to career changes, "author name," "knowing about an event in advance," and "purchase history" from new hobbies and activities.

The experiment revealed that the groups of variables extracted as unexplored data differed among the three topics. Participants often made comments in Topics 2 and 6 that centered on themselves, thereby reflecting changes in their mental states. Therefore, such topics can be characterized by having many variables that originate from psychological changes, including attitudes. Meanwhile, for Topic 5, participants often mentioned changes in other people; consequently, easily quantifiable variables such as "annual salary," "sales," and "labor hours" appeared more than variables stemming from mental states.

Moreover, the number of variables obtained by entering the comments for each topic into the VQ were 24, 25, and 27, respectively. However, the number of cases for all three topics in which these variables were included as words in the set of comments collected by the MROCs was zero. In other words, the keywords representing the variables were not included in any of the words spoken by the subjects. These variables are from past data such as the "Sense of Crisis Database," "Suicide Statistics Data," or "Event and Shopping Mall Sales Data" and are thus difficult to apply directly on ongoing events such as the COVID-19 pandemic. Nonetheless, decomposing the past data in variable units and reconstructing them using VQ could contribute to unexplored data discovery, as a set of variables from texts not containing any information related to the data.

Acknowledgments

This study was supported by NEC Corporation. We would like to thank PLUG-Inc. in the survey design of MROCs.

References

- Baldus, B.J. 2015. Insight Generation with Marketing Research Online Communities (MROCs). *J. Internet Commer.*, 14:476–491.
- Hayashi, T., Ohsawa, Y. 2017. Matrix-based Method for Inferring Variable Labels Using Outlines of Data in Data Jackets. *PAKDD*.
- Hayashi, T., Ohsawa, Y. 2020. Data Origination: Human-centered Approach for Design, Acquisition, and Utilization of Data, *Int'l Conf. Soft Computing and Pattern Recognit.*