

An Analysis of Chronological Changes Based on Feature Words Using the Cross-Tabulation Search Engine - Case Study for Self-Assessment Documents in Japanese University -

Kahori Ogashiwa ^{*}, Kumiko Kanekawa [†], Toru Sugihara [‡],
Soichiro Aihara [§], Masao Mori ^{**}, Sachio Hirokawa ^{††}

Abstract

In Japan, university evaluation is compulsory, and many universities issue self-assessment reports yearly, often with the same table of contents. With this in mind, feature words from each chapter can extract for documents with nearly identical hierarchical structures. For example, by setting “table of contents” on the vertical axis and “fiscal year” on the horizontal axis and visualizing the feature words in a cross table, it is possible to find chronological changes in the feature words at a glance. This makes it possible to quickly grasp what has changed or not in each chapter. The present study looked at the chronological changes in Yamaguchi University's self-assessment documents from 2012 to 2016. With the developed system, we discovered a significant chronological change in Yamaguchi University's Diploma Policy, which is one of three policies. Yamaguchi University had already formulated and operated a similar policy called the Graduation Policy before the Diploma Policy became mandatory. The system showed that Yamaguchi University changed the “Graduation Policy” to “Diploma Policy” in line with the mandatory Diploma Policy. By using this system, even those who are not affiliated with Yamaguchi University have been able to grasp the significant changes in the past at the University. The reasons for any change in university policy should be explained to its stakeholders. The system can be an effective means for examining the decision-making process.

Keywords: chronological change, cross-tabulation, feature word, self-assessment document

1 Introduction

So far, we have analyzed a large number of self-assessment documents consisting of several hundred pages with human eyes regarding university evaluation in Japan [1]. We have checked the feature words and conducted university analysis and inter-university comparisons. We have developed a system that utilizes text mining technology to reduce this enormous amount of labor cost [2]. We were able to greatly improve efficiency and obtain new knowledge that could not be discovered by human eyes without a massive amount of work by using the developed system

^{*} Ochanomizu University, Tokyo, Japan

[†] Research Organization of Information and Systems, Tokyo, Japan

[‡] Kanto Gakuin University, Yokohama, Japan

[§] Shibaura Institute of Technology, Tokyo, Japan

^{**} Tokyo Institute of Technology, Tokyo, Japan

^{††} Advanced Institute of Industrial Technology, Tokyo, Japan

“Cross-Tabulation Search Engine” [2]. By utilizing the cross-tabulation search engine and applying it to university self-assessment documents, it is possible to efficiently set hypotheses about chronological changes without reading a huge number of documents. The capability of this system can clarify by verifying the case study results. A cross-table is created in this system based on the table of contents structure of the document and the publication year of the document, allowing for efficient hypothesizing of chronological changes. In the present study, we report a case in which we were able to set hypotheses efficiently and accurately about chronological changes by applying the method to Yamaguchi University's self-assessment reports. The authors read the actual self-assessment documents verifying that the hypothesis is correct.

2 Related Work

In university evaluation, yearly comparisons are performed using numerical data using various indices [3]. Visualization of numerical data related to year-to-year comparisons (comparisons of chronological changes) is commonly performed in the fields of university evaluation and institutional research [4][5]. Yamaguchi University, which is the subject of the present study, has also developed a self-assessment system. The system displays graphs of numerical data comparisons over time, and evaluators have created self-assessment reports using these graphs [6]. Recently, there has been a lot of research on text mining in university evaluation. We have researched practical research on data-based university evaluation. The purpose of the above-mentioned study was to focus on various documents (text data) expressing university activities and to develop a system that visualizes the analysis results as a two-dimensional map using text mining technology. National and public universities are required to disclose their medium-term plan documents to all universities. As a result, despite the slightly different format of the document, the result of analyzing the two-dimensional map when searching for the characteristics of national universities with the keyword “region (community)” has been confirmed to be a valid result [2][7][8]. In addition, two-dimensional visualization of public universities has yielded useful results [9][10].

As far as the authors are aware, no studies have been conducted on self-assessment documents comparison of chronological changes. As a result, in this study, we extract feature words from self-assessment documents using text mining and visualize these feature words as a cross table of the table of contents and document publication year to detect a change in feature words.

3 Cross-Tabulation Search Engine

The cross-tabulation search engine used in the present study is one of the many systems developed by Professor Hirokawa [11]. As a previous study directly related to the present study, there is a cross-tabulation search engine for the medium-term plan, which was mentioned in “2. Related work” [2]. As shown in Figure 1, typing “z” in the box will display all search results. By setting “view” on the vertical axis and “year” on the horizontal axis, it is possible to freely set how many top-level feature words to display. In Figure. 1, it is an example of setting “v” to “view”, “h” to “year”, “feature” to “5”. See reference [2] for details of the cross-tabulation search engine.

The cross-tabulation search engine used in this study has recently added two functions: a feature word English translation function, and a function to detect non-prototype representations when the feature word is a verb (Google automatic translation). For example, if there is a verb such as “求める (in Japanese)” in the feature word that should be displayed in color, it may not

be the original form like “求め (in Japanese)”. In this system, when the search results are displayed in the lower frame, morphological analysis is performed, and if the feature word is a verb, occurrences other than the base form of the verb are also colored (Figure 2). If you check “en” in the menu and click the “Search” button, the feature words and the words on the vertical and horizontal axes will be displayed in English (Figure. 1). A number before the feature word denotes the frequency of occurrence.

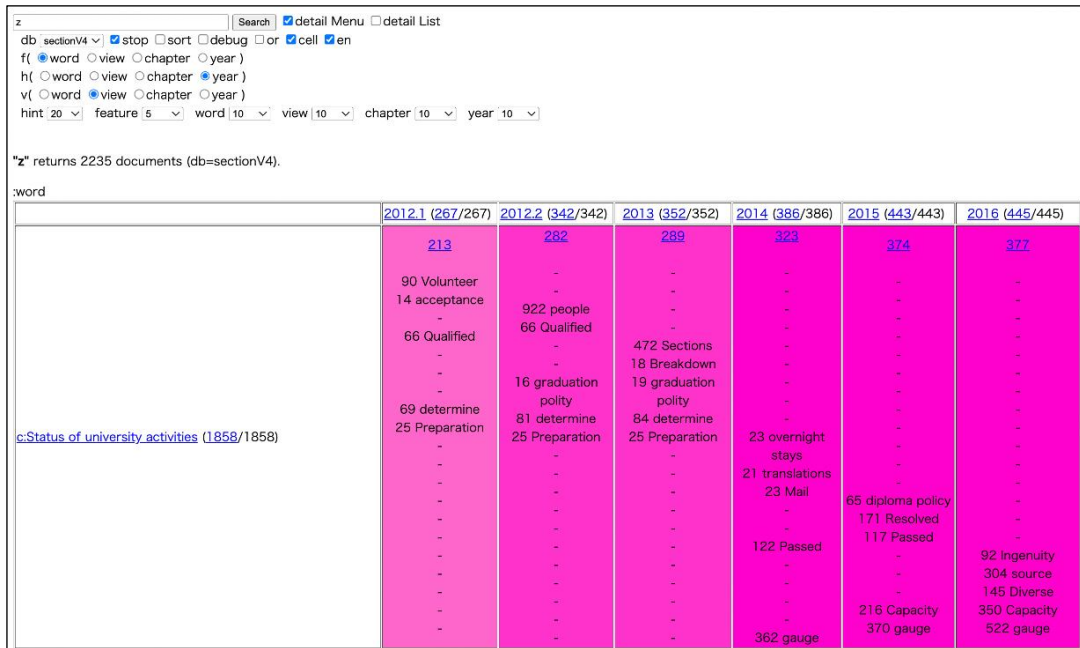


Figure 1: Cross-tabulation search engine setting screen and search results



Figure 2: The color display of feature words and feature words in sentences (bottom frame). For the lower frame, even if the feature word is the prototype, the non-prototype is also colored. For example, if the feature word is “求める,” “求め” other than the prototype is also colored (green in this example).

4 Case Study

In this chapter, we report the results of analyzing the relationship between Yamaguchi University's self-assessment report [12] and trends in educational policy in Japan. The reason for using this case is that the author was working at Yamaguchi University at the time, which could render the author more able to judge the validity of the search results of the system. All Japanese universities are required to create and publish three policies—namely, admission policy, curriculum policy, and diploma policy, as of April 2017. We examined changes at Yamaguchi University during the same time period with a focus on this point. By using the developed system, it became clear that there was a change (differences of feature words) in the descriptions of the admission policy and the diploma policy. Feature words related to curriculum policy were omitted due to space limitations.

4.1 Admission Policy

The study hypothesizes that stricter quota management for private university entrance examinations, which began in fiscal year 2016, will affect the admission policy. Figure 3 shows the search results of this system for admission policies. Compared to 2012 to 2015, in 2016 new feature words are appearing. Specifically, the words [selection (選抜)] and [magnification (倍率)] appeared as new feature words in 2016. We found there had been a change in Yamaguchi University.

	15	17	17	18	19	19
	25 want	25 want	24 want	-	-	-
	17 interest	18 interest	18 interest	27 interest	-	-
	36 seek	36 seek	37 seek	53 seek	-	-
	-	-	-	-	30 leave	-
	-	-	-	-	13 Reading	-
	-	-	-	36 Academic	comprehension	-
	23 Willingness	23 Willingness	25 Willingness	38 Willingness	39 Academic ability	-
	-	-	-	-	38 Willingness	-
	-	-	-	-	-	13 AP
	-	-	-	-	-	64 selection
	-	-	-	-	-	17 times
	-	-	-	-	-	23 magnification
	20 like	20 like	21 like	27 like	29 like	35 like

Figure 3: Admission policy search results

4.2 Diploma Policy

Figure 4 shows the search results of this system for diploma policies. Although [year] is set on the horizontal axis, nothing is displayed in the three columns on the left (1 in 2012, 2 in 2012, 2013), and the table of contents of the self-assessment reports in 2013 and 2014 has changed. The Graduation Policy was being used by Yamaguchi University as an alternative to the Diploma Policy up until 2013, according to another document (on the university's website), which we used to confirm this change. Yamaguchi University's Graduation Policy has been renamed the diploma policy as a result of the government's decision to create and publish the three policies (because it was made mandatory). We were able to verify this from the corresponding part of the text of the self-assessment documents. Figure 4 shows the results of the search for the Graduation Policy, but the three columns on the right (2014–2016) are empty because they have changed to the Diploma Policy.

To verify the hypotheses of changes discovered by using this system, the text of the self-assessment reports was checked by human eyes, and it was found that it had changed since 2014, indicating the usefulness of this system. The Graduation Policy has been renamed to the Diploma Policy, but only Yamaguchi University members and stakeholders are aware of this change. However, because this system makes the change visually apparent, even people outside of Yamaguchi University can immediately notice it.

	85	98	97	103	110	111
	15 Way of thinking	-	-	-	-	-
	14 Explore	23 acceptance	-	-	-	-
	-	-	28 rich	-	-	-
	27 rich	11 graduation polity	14 graduation polity	-	-	-
	-	58 Policy	57 Policy	-	-	-
c:Basic policy of the curriculum and its outline (604/604)	58 determine	63 determine	65 determine	174 put on	63 Thinking	-
	-	-	-	207	73 Thinking	-
	-	-	-	Knowledge	250 Knowledge	111
	-	-	-	-	-	Communication
	-	-	-	-	-	68 Adopted
	-	-	-	-	-	173 source
	-	-	-	39 diploma policy	49 diploma policy	54 diploma policy
	34 policies	36 policies	38 policies	55 policies	65 policies	73 policies

Figure 4: Diploma policy search results

5 Conclusion

In the present study, the cross-tabulation search engine we have created so far can be used to efficiently set hypotheses about chronological changes without having to read a ton of materials by applying it to university self-assessment documents. We were able to verify this from the results of the case analysis of Yamaguchi University. By using this system, we were able to verify that hypotheses about [past changes] can be efficiently formulated, in other words points of change can be discovered, from changes in feature words by year. Additionally, based on the findings of the case study, it became possible for people who are not connected to Yamaguchi University to use this system to comprehend past changes at Yamaguchi University, in other words changes in feature words. The reasons for any change in university policy should be explained to stakeholders of the university. The system can be used as an effective method for examining the decision-making process of universities. It was suggested that this system could be used to piece together records of previous processes. The present study targeted self-assessment documents with the same structure of table of contents. In the future, we plan to apply to documents with different table of contents structures.

Acknowledgement

This work was supported by JSPS KAKENHI Grant Number JP20K02975, JP21K02659.

References

- [1] National Institution for Academic Degrees and Quality Enhancement of Higher Education (NIAD-QE), Glossary of Quality Assurance in Japanese Higher Education 5th edition, 202

- 1, https://niadqe.jp/glossary/#term_en (accessed 2022-10-12).
- [2] Kahori Ogashiwa, Masao Mori, and Sachio Hirokawa, “Text Mining of the Midterm Goals of National Universities for the Development of their Individual Characteristics,” *The Journal of Information and Systems in Education*, vol.19, no.1, pp.15-20, 2020, <https://doi.org/10.12937/ejsise.19.15> (accessed 2022-10-12).
- [3] IR Office, Sophia School Corporation, Sophia Facts, 1-1 Sophia University Education (Undergraduate), <https://dept.sophia.ac.jp/ir/en/1-1/#1-1-1> (accessed 2022-10-28).
- [4] Tetsuya Oishi, “What is the Essential Curriculum for IR in Japan?, ” *IIAI Letters on Institutional Research*, vol.1, 2022, <https://doi.org/10.52731/lir.v001.009> (accessed 2022-10-30).
- [5] Noriko Ito, Nobuhiko Seki, Masao Mori, “Changes in the Social Positioning of Higher Education and the Emergence of New Great Divergence,” *IIAI Letters on Institutional Research*, vol.1, 2022, <https://doi.org/10.52731/lir.v001.042> (accessed 2022-10-30).
- [6] Kahori Ogashiwa, Takahiro Matsumoto, Yue Wang, Joji Kariya, Hiroo Naitoh, “Evaluation of the Yamaguchi University Self-Assessment and Evaluation System and Its Improvement,” *International Journal of Institutional Research and Management*, vol.3, no.1, pp.1-14, 2019.
- [7] Kahori Ogashiwa, Eiichi Takata, Tetsuya Oishi, Masao Mori, Sachio Hirokawa, “Automatic Estimation and Feature Word Analysis of Universities Using University Medium-term Plan,” *Proceedings of the International Congress on Advanced Applied Informatics (IEEE/IIAI-AIT 2019)*, Yogyakarta, Indonesia, pp.39-44, November 2019.
- [8] Kahori Ogashiwa, Eiichi Takata, Tetsuya Oishi, Masao Mori, Sachio Hirokawa, “Text Mining and Logic Model for University Midterm Plans,” *Proceedings of the 8th International Congress on Advanced Applied Informatics (IIAI-AAI 2019)*, Toyama, Japan, pp.1013-1014, July 2019.
- [9] Soichiro Aihara, Toru Sugihara, Kahori Ogashiwa, Kumiko Kanekawa, Masao Mori, Sachio Hirokawa, “A Preliminary Study on the Medium-term Plan of Public Universities Transferred from Private Universities,” *Proceedings of the 11th International Congress on Advanced Applied Informatics (IIAI-AAI2021 Winter)*, Online, December 2021.
- [10] Soichiro Aihara, Kahori Ogashiwa, Masao Mori, Sachio Hirokawa, Kumiko Kanekawa, Toru Sugihara, “Text Analysis to the Preambles of the 4th Medium-term Goals/Plans of National University Corporations,” *IIAI Letters on Institutional Research*, vol.1, 2022, <https://doi.org/10.52731/lir.v001.024> (accessed 2022-10-30).
- [11] Sachio Hirokawa, Eisuke Ito, Kensuke Baba, “Visualization of Keywords for Exploratory Search of Research,” *Journal of Information Processing and Management*, vol. 58, no. 6, pp. 447–454, 2015, <https://doi.org/10.1241/johokanri.58.447> (in Japanese).
- [12] Yamaguchi University self-assessment and evaluation, <http://committee.ue.yamaguchi-u.ac.jp/jikotennkennhyoka-system.html> (accessed 2022-10-30).