웹사이트에 게시된 기업의 소개글 분석을 통한 기업의 현재 및 미래 가치 예측 분석 방법

The Association of Institutional Information on Websites with Present and Future Financial Performance

나형종(Hyung Jong Na)[*], 최석재(Sukjae Choi)[**], 권오병(Ohbyung Kwon)[***]

초 록

기업의 "회사소개" 페이지는 기업이 중요시하게 여기는 가치가 무엇인지를 소개하고 있다. 본 연구에서는 이에 관한 글이 회사의 현재 및 미래의 재무성과의 연관성을 조사한다. 텍스트 마이닝 기법을 활용하여 2016년의 S&P500 기업의 기관 정보를 분석하여 회사소개 페이지에서 의도적으로 반복되는 단어를 파악하고, 이를 통해 요인분석을 수행하였다. 그 결과 홈페이지의 키워드는 6가지 요인으로 분류될 수 있음을 알 수 있었고, 그후 최소 자승 회귀분석을 수행하여 키워드의 요인과 기업의 현재 재무 성과 사이의 연관성을 결정하였다. 분석 결과 요인 2("구매 경험"과 관련된 키워드)가 현재 재무 성과의 대체 변수인 ROE와 긍정적인 상관관계가 있음을 보여주며, 요인 1("고객에게 알림"과 관련된 키워드)은 ROE와 부정적인 관계를 가짐을 알 수 있었다. 다만, 요인 1은 기업의 미래 재무 성과를 측정하는 대안 변수인 Tobin's Q와 긍정적인 관계가 있었다. 이 결과는 기업의 웹사이트에 소개된 글과 기업의 재무 성과 사이에 관계가 있음을 나타낸다. 따라서 웹 사이트에 대한 기관의 소개글은 현재 회사 성과 및 미래 회사 가치에 대하여 유용한 지표가 될 수 있다.

ABSTRACT

The "About Us" page on the website of a corporation provides information regarding the organization's vision, philosophy, and values. We examine the association between institutional information provided on corporate websites (i.e., the "About Us" section) with present and future financial performance. Utilizing a text mining technique, we analyze the institutional information of S&P500 firms in the year 2016. We conduct a factor analysis including words that are intentionally repeated in the introductory text of corporate websites. The results of the analysis reveal that keywords from this institutional information can be grouped into six factors. We then carry

out an ordinary least squares regression analysis to determine the associations between these six factors and present financial performance. The results show that keywords in Factor 2 (those related to *Purchasing experience*) are positively associated with ROE, a variable representing present financial performance, while keywords in Factor 1 (those related to *Note to customers*) show a negative relationship with ROE. On the other hand, keywords in Factor 1 have a positive relationship with Tobin's Q, a variable representing future financial performance. These results indicate that there is some relationship between the words used in the institutional information in this section of corporate websites and firms' financial performance. Hence, the institutional information on a website may be a useful indicator of current firm performance and future firm value.

1. Introduction

An introductory message is often provided on corporate websites to summarize information about the company. This information may concern the company's vision, future strategies, and corporate culture [35]. On corporate websites, these "About Us" sections form a first impression in the reader's mind about the corporation [39]. A company also communicates with many groups through its website. Leporini and Paternò [26] explain that firm performance can be increased by improving website usability in order to pursue the firm's goal. Information about current and future financial performance and the firm's business model may be obtained from the "About Us" section.

We empirically investigate how financial information that can aid in present or future valuation of a firm can be inferred from institutional information on corporate websites. Estimating financial performance using this institutional information, which is freely accessible at any time and any place, may be much easier and more effective than relying on periodically published financial reports. We examine institutional information as a useful indicator of firms' current financial performance and future firm value.

Meanwhile, text mining techniques become more useful tools to extract meaningful determinants of the firm's performance [6, 19, 22]. Utilizing a text mining technique, we analyze institutional information in the "About Us" sections of websites of S&P500 firms in the year 2016. Through a factor analysis of repeated words, we identify six factors. An ordinary

least squares (OLS) regression analysis enables us to determine the association between these factors and firm performance. The results show that keywords in Factor 2 (those related to *Purchasing experience*) have a positive relationship with ROE (return on equity), a variable representing present financial performance. From these data, we infer that a corporate focus on customer convenience and satisfaction is associated with relatively high financial performance. The results of the OLS regression analysis also show that keywords in Factor 1 (those related to *Note to customers*) have a negative relationship with ROE. This result implies an association between prioritizing product information on the website and relatively lower current performance. On the other hand, keywords in Factor 1 (those related to *Note to customers*) have a positive relationship with *Tobin's Q*, a variable measuring future financial performance. We find that firms that prioritize product information in the "About Us" section have relatively high future corporate value. That is, we find an association between active communication with customers and positive business outcomes, despite the expense of webpage maintenance.

The contribution of this paper is as follows. First, to the best of our knowledge, this is the first study to propose a method of analyzing non-financial, unstructured institutional information on corporate websites combining a text mining method and OLS to infer financial performance. Second, this empirical study

demonstrates an apparent relationship between information design on company websites, especially in terms of content, and financial performance. Here, A website's information design refers to the site's ability to deliver relevant, current, and easy-to-understand information to its users [15]. We conclude that institutional information on firms' websites can be used to infer financial performance, including current and future corporate value.

The remainder of this paper is organized as follows: in section 2, we provide a review of literature on institutional information. Our data collection method, including text mining, and our research model, including the independent variables identified by text mining, are described in section 3. The results, corresponding implications, and conclusion are provided in sections 4 and 5, respectively.

2. Institutional Information on Corporate Websites

Companies use their websites as marketing tools to provide potential customers with information [2]. Firms may also communicate with consumers through their websites [29]. Marketing through websites can be crucial for successful performance [13] and the firm‐customer relationship. Companies can share their corporate vision and goals with their customers through their websites; having a shared vision and building trust enhances the firm‐

customer relationship, which positively affects customer retention [26, 27, 32].

A corporate website is a company-controlled communication tool. Feedback opportunities give the company information regarding the needs of specific customers, which can be used to customize products and services [8]. To prevent serious damage to customer relationships, managers must monitor communications carefully, analyze customer feedback, and react appropriately.

Corporate websites can also facilitate inter-organizational communication. Strategic dialogue conducted via company webpages (including SNS websites [20]) can help to establish and foster relationships with the public [9, 42]. Moon [31] studied corporate public relations development programs utilizing website communications, evaluating their effectiveness.

In addition, the introductory text on a company's website is located in a visually conspicuous place (the "About Us" section) that is easily accessible to users and includes important information about the company. Heterogeneity of informational design in the introductory text may improve competitive advantage over other companies. However, studies on website content, especially the introductory page and the institutional information in the "About Us" section, and its effect on firm performance have been lacking. In fact, the association between information design and firm performance has not been em-pirically tested, despite its possible important implications.
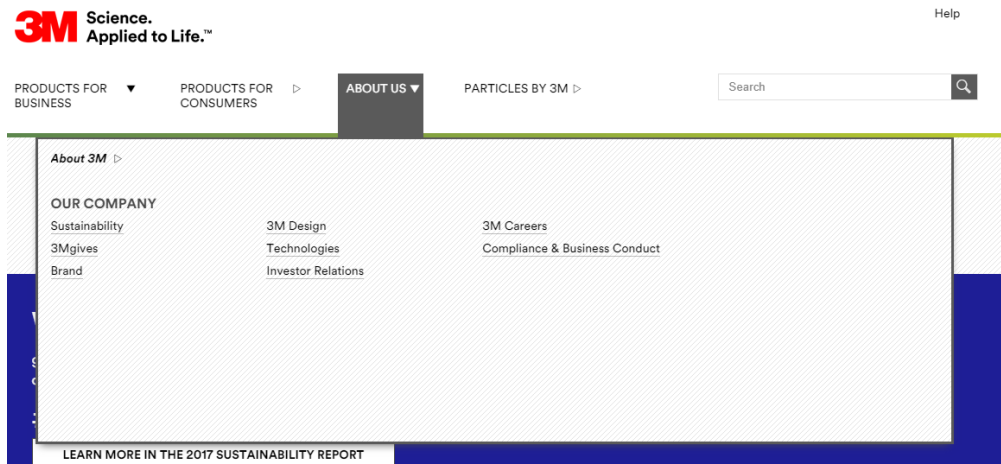
3. Methods

3.1 Data Collection via Text Mining

We conducted an empirical analysis regarding the association between the introductory text in the "About Us" section on firms' websites and financial performance. The sample includes S&P500 companies in the year 2016. To increase the reliability of the analysis, we chose companies among S&P500 firms which satisfy the following conditions. Financial data used in this study were collected from the Osiris Data Base (https://osiris.bvdinfo.com/) by Bureau van Dijk. Firms included in the sample satisfied the following criteria:

(1) S&P500 listed companies in 2016
(2) Companies from non-financial industries whose fiscal year-end was in December
(3) Firms for which required financial data were available in the Osiris system.

Introductory information is typically located in the "About Us" section of a company's website. We used the Java function to identify pages containing firms' introductory messages. Keywords were extracted from the institutional information (the "About Us" section) on the websites of S&P500 companies using a text

〈Figure 1〉 3M's "About Us" menu

mining method, and the term frequency‐inverse document frequency (TF–IDF) was obtained for each keyword. We now outline this process in more detail for an example company, 3M.

On the 3M homepage, in the "About Us" menu, there are eight headings: Sustainability, 3 Mgives, Brand, 3M Design, Technologies, Investor Relations, 3M Careers, and Compliance & Business Conduct, as shown in <Figure 1>. Under each of these headings is basic information about the firm. In this study, we included all relevant information from these headings in our analysis. Any subheadings were also included in the analysis. For instance, for the 3M company, the Sustainability heading has the following subheadings: Our Every Life Ambition, Our Purpose-Driven Business, Global Challenges, Industry Solutions, and Goals and Progress. These also contain information about our main area of focus: the firm's vision, shared

values, and corporate culture. However, we excluded any text related to financial statements, company history, award records, geographic location, contact information, policies, introduction of board members, and partner companies, and anything else unrelated to our main area of focus (the firm's vision, shared values, and corporate culture).

Some companies have different names for the "About Us" section of their websites, such as "About Our Company", "Who We Are", "We Are [company name]", or "Discover [company name]". One example is Adobe, for which the introductory text can be found under the heading, "About". Other subheadings include Vision, Mission, Principles, Philosophy, Beliefs, and Sustainability. These and other subheadings containing introductory text were included in the analysis. In some cases, as many as 31 pages for a given firm included introductory text relevant for our purposes.

For firms lacking an official introductory text, we analyzed the text on the default page of the homepage, which usually showcases the firm's spirit and values. For instance, Alphabet, Google's mother company, does not have an "About Us" page. However, there is a summary of Google on the first page. We listed companies by name and then recorded the website addresses on which the introductory text was located in an Excel sheet, as shown in <Figure 2>.

After collecting the text and recording the website addresses, we created a single computer file for each company by combining all text from the introductory messages written on all relevant pages on each company's website. Companies with no website or with websites under construction were excluded from the analysis. In total, introductory text was collected for 354 companies.

We identified the radix of the words in each of the saved files. English is a language in

Company	Address 1	Address 2	Address 3	Address 4
3M Company	http://www.3m.com/3M/en_US/sustainability-us/	http://www.3m.com/3M/en_US/sustainability-us/every-life/	http://www.3m.com/3M/en_US/sustainability-us/global-challenges/	http://www.3m.com/3M/en_US/sustainability-us/industry-solutions/health-care/
AES Corporation	http://www.aes.com/about-us/about-us-overview/default.aspx	http://www.aes.com/about-us/mission-vision-values-and-culture/default.aspx	http://www.aes.com/about-us/ethics-and-compliance/default.aspx	http://www.aes.com/about-us/our-history/default.aspx
Abbott Laboratories	http://www.abbott.com/about-abbott/abbott-citizenship/finding-the-upside.html	http://www.abbott.com/about-abbott/abbott-citizenship/upside-in-action.html	http://www.abbott.com/about-abbott/abbott-citizenship/citizenship-reporting.html	http://www.abbott.com/about-abbott/who-we-are.html
Abbvie Inc	http://www.abbvie.com/about-us/home.html	http://www.abbvie.com/about-us/who-we-are.html	http://www.abbvie.com/about-us/ethics-compliance.html	
Accenture PLC	https://www.accenture.com/kr-en/insight-data-ethics			
Activision Blizzard	http://www.activisionblizzard.com/about-us			
Acuity Brands	http://www.acuitybrands.com/about-us	http://www.acuitybrands.com/about-us/careers		
Adobe Systems INC	https://www.adobe.com/about-adobe.html	https://www.adobe.com/about-adobe/explorations.html	http://www.adobe.com/corporate-responsibility/education.html	http://kickbox.adobe.com/
Advance Auto Parts Inc	http://corp.advanceautoparts.com/about/general.asp	http://corp.advanceautoparts.com/about/public.asp		

〈Figure 2〉 URLs Containing Institutional Information about S&P500 Companies

which word form changes as grammatical content changes. For example, the words *product* and *products* only differ in number (i.e., whether they are singular or plural). They have virtually the same meaning otherwise. In this study, we are only interested in the meaning of each word. Word form changes are irrelevant; therefore, we split the radix of each word in order to exclude variations in word form using Porter's [34] algorithm and coding in Java. We removed any non-words. For example, a basic HTML tag was removed using the text extraction process that is a typical Java function. However, not all symbols presented obstacles to the analysis. Thus, we excluded punctuation, but included symbols (e.g., ∽+−=±÷♥♡★☆ ♪♬) that have no linguistic meaning but may communicate something to customers. We then calculated the TF-IDF value of each word. Since it is difficult to categorize documents by word frequency only, we obtained TF-IDF values. The IDF is the inverse value of the number of times a word appears in a document. For example, if word A appears ten times in

five hundred documents, the IDF is 500 divided by 10, which is 50. Usually, the IDF value is large. Therefore, we used the log value of IDF [5]. In each document file, we chose keywords with frequencies larger than 1%. Hence, not all words were present in all files. The purpose of this procedure was to identify unique keywords expressing the traits of the company. Then, we rearranged the words used in each file and recorded their frequencies to avoid any redundancy. Additionally, we recorded how many times these words were used in each file. TF-IDF values may vary despite the fact that the same word appears multiple times. After recording TF-IDF values for each file, we saved these data in .csv files.

Next, we linked financial information about S&P500 companies included in the analysis to information in the .csv file. Financial information for these companies and TF-IDF values for each company were linked by company name and saved in a separate file. The results are shown in <Figure 3> below.

Company names are listed in the far left

Company Name	Financial information							TF-IDF			
	SIZE	LEV	GRW	PPE	TQ	ROA	ROE	food	us	....	peltor
3M Company	24.2162	0.65243	−0.00545	0.03165	0.65243	0.1535	0.44185	223	370	....	200
Abbott Laboratories	24.4428	0.48575	0.02195	0.06131	0.48575	0.0339	0.06600	0	0	....	0
Abbvie Inc	24.6945	0.92563	0.12157	0.02327	0.92563	0.1122	1.50899	0	0	....	0
Accenture PLC	23.6283	0.63606	0.05721	0.00264	0.63607	0.2251	0.61855	0	20	....	0
Activision Blizzard	23.4475	0.47081	0.41680	0.01764	0.47081	0.0633	0.11973	0	0	....	0
Acuity Brands	21.601	0.43498	0.21598	0.01641	0.43498	0.1208	0.21382	0	22	....	0
Adobe Systems INC	23.1851	0.40292	0.22081	0.00602	0.40292	0.0996	0.16693	0	24	....	0
Advance Auto Parts Inc	22.8185	0.69725	−0.01739	0.01315	0.69725	0.0565	0.18678	0	14	....	0
AES Corporation	24.3197	0.91365	−0.04019	0.53369	0.91365	−0.0312	−0.36233	0	55	....	0

〈Figure 3〉 TF-IDF Values and Financial Information

column. Each firm's financial information is recorded from the second column to the seventeenth column in each row. Companies' TF-IDF information for 1,099 words is recorded on the far right side. As we can see in <Figure 3> above, many words appear in certain companies' documents but not in others, making it difficult to obtain meaning, since it was important to the success of this analysis to find companies with similar traits. Deleting columns containing unnecessary words minimizes the number of dimensions, allowing for a more appropriate analytical process and facilitating understanding. Therefore, columns containing words that seldom appear were deleted. After setting the criteria to a minimum of two appearances, we were left with 286 columns. This allowed us to delete 70% of the initially identified dimensions. A minimum of 15 appearances of a keyword was required for its inclusion in the analysis.

Finally, 304 firms for which financial information and introductory text were available were used in the analysis. To identify associations of various keywords in the institutional information provided on corporate websites with current company performance and future corporate value, we extracted six factors by factor analysis. The process of factor analysis is as follows. First, when a word is found in the website, we specified 1 as a dummy variable, otherwise 2. Nonsense words (e.g., ai, ga, etc.), words related to specific industries (e.g., aerospec, automobil, medicin, etc.) or nations (e.g., english, fran, espa, etc.), and words related to homepage functions (e.g., toggle, navi, etc.) were excluded. Next, factors with eigenvalues greater than 1.2 were chosen to determine factor numbers. By this process, the sample was divided into six groups. We then used the Varimax method to choose keywords included in these six groups. Additionally, to test the internal validity of each factor, we excluded keywords with Cronbach's alpha (α) coefficient values of no more than 1.2. Through this process, we narrowed down the list of keywords in each group. Each factor was named with consideration of the common characteristics of the

〈Table 1〉 Firms' Introduction Words in Their Homepages

Factor	Keywords				Name
Factor 1	safeti	newsroom	overview	devic	*Note to Consumers*
Factor 2	shop	store	card		*Purchasing Experience*
Factor 3	trademark	logo			*Corporate Identity*
Factor 4	cloud	network	internet		*Network Infrastructure*
Factor 5	storag	payment	supplier		*Supply Chain*
Factor 6	copyright	blog	specialti	info	*Misc. Information*

Note) since text mining analysis classifies words into root of word, above words in the table such as "safeti", "devic", "storag", "specialti", "info" result in as a keyword.

keywords in terms of meaning. Six factor groups of keywords were established: *Note to customers* (Factor 1), *Purchasing experience* (Factor 2), *Corporate identity* (Factor 3), *Network infrastructure* (Factor 4), *Supply chain* (Factor 5), and *Misc. information* (Factor 6). Six factor groups of keywords are shown in <Table 1> as below. Note that we did not find a meaningful set of words from LDA algorithm.

3.2 Research Model

We divided firms' financial performance into two categories: current financial status and future corporate value [33]. We used *ROE* as a proxy for present firm performance according to the literature [4, 24, 44]. *ROE* is calculated as net income in the current term divided by the sum of equity. We also used *Tobin's Q* as an alternative variable representing future corporate value [11, 40, 41]. *Tobin's Q* is calculated as the sum of the market value of equity and the book value of debt divided by the book value of total assets. Market value means stock price. This implies future expected income.

To test the association between keywords in introductory text and financial performance, we established two models as shown in Equations (1) and (2). The independent variables are our six factors: *Note to customers, Purchasing experience, Corporate identity, Network infrastructure, Supply chain*, and *Misc. information*.

Control variables were also included in the research model. First, firm size (*SIZE*) is considered because the larger the firm, the higher its value and the greater the opportunity for growth [12]. At the same time, company size may also have a negative effect on firm performance because of high political costs. Second, the debt ratio (*LEV*) is included due to the evidence that higher firm performance is associated with a higher debt ratio [36]. A higher debt ratio can increase the extent of monitoring in the capital market, which may also be relevant to financial performance. Therefore, *LEV* is expected to be negatively associated with firm performance. Third, sales growth (*GRW*), which indicates the amount of increase in this year's sales compared to last year's, is used to control for corporate growth that might influence corporate value [28]. Next, *PPE* (property, plant, and equipment) is related to firm performance. As the proportion of tangible assets grows, depreciation and maintenance costs grow as well. This can impact corporate outcomes. Possession of large amounts of tangible assets may have a positive influence on firm value and performance [43]. Finally, to control for industry effects, we employ an industry dummy in the OLS model. Using Equations (1) and (2), we then perform the OLS regression analysis as follows.

If β_1, the coefficient value of the independent variables for Factors 1 to 6, shows a positive and significant result, this means that if the introductory text on a corporate website con-

tains more keywords related to these variables, current management performance will be comparatively better or future corporate value will be higher. However, if β_1 shows a negative and significant result, if the institutional information contains more of these keywords, current management performance will be comparatively poorer or future corporate value will be lower.

$$
\begin{aligned}
ROE_t = &\ \alpha_0 + \beta_1 Eactor\,1_t (Factor\,2_t \qquad (1)\\
&/Factor\,3_t/Factor\,4_t/Factor\,5_t\\
&/Factor\,6_t) + \beta_2 SIZE_{t-1} + \beta_3 LEV_{t-1}\\
&+ \beta_4 GRW_s + B_5 OCF_t + \beta_6 PPE_t\\
&+ \beta_7 Industrry_t
\end{aligned}
$$

ROE_t = net income of the current term, all divided by the sum of equity

$Factor1_t$ = the average of TF-IDF of the keywords group related to Note to customers

$Factor\,2_t$ = the average of TF-IDF of the keywords group related to Purchasing experience

$Factor\,3_t$ = the average of TF-IDF of the keywords group related to Corporate identity

$Factor\,4_t$ = the average of TF-IDF of the keywords group related to Network infrastructure

$Factor\,5t$ = the average of TF-IDF of the keywords group related to Supply chain

$Factor\,6t$ = the average of TF-IDF of the keywords group related to

Misc. information

$SIZE_{t-1}$ = natural logarithm result of total assets

LEV_{t-1} = debt ratio (= total debt/total capital)

GRW_t = (sales of this year/sales of last year)-1

PPE_{t-1} = property, plant, and equipment.

$$
\begin{aligned}
TQ_t = &\ \alpha_0 + \beta_1 Factor\,1_t (Factor\,2_t \qquad (1)\\
&/Factor\,3_t/Factor\,4_t/Factor\,5_t\\
&/Factor\,6_t) + \beta_2 SIZE_{t-1} + \beta_3 LEV_{t-1}\\
&+ \beta_4 GRW_t + \beta_5 OCF_t + \beta_6 PPE_t\\
&+ \beta_7 Industry_t
\end{aligned}
$$

TQ_t = the sum of the market value of equity and the book value of debt, all divided by the book value of total assets

$Factor\,1_t$ = the average of TF-IDF of the keywords group related to Note to customers

$Factor\,2_t$ = the average of TF-IDF of the keywords group related to Purchasing experience

$Factor\,3_t$ = the average of TF-IDF of the keywords group related to Corporate identity

$Factor\,4_t$ = the average of TF-IDF of the keywords group related to Network infrastructure

$Factor\,5_t$ = the average of TF-IDF of the keywords group related to Supply chain

$Factor\,6_t$ = the average of TF-IDF of the keywords group related to Misc. information

$SIZE_{t-1}$ = natural logarithm result of total assets

LEV_{t-1} = debt ratio (= total debt/total capital)

GRW_t = (sales of this year/sales of last year)−1

PPE_t = property, plant, and equipment

$Industry_t$ = dummy variable to control for industry effects.

4. Results

Descriptive statistics related to the six factors, financial variables, and other variables used in this study are shown in <Table 2>. The final sample included 304 firm-year observations. The average value of *Tobin's Q (TQ)*, a variable representing future firm value, is 2.359. The median of *TQ* is 1.970 and the standard deviation of *TQ* is 1.278. The average value of *ROE*, a variable representing present firm performance, is 0.120. The median of *ROE* is 0.158 and the standard deviation of *ROE* is 0.618. Variables for Factors 1~6 are the average of TF-IDF dummy variables for each keyword.

〈Table 2〉 Descriptive Statistics (n=304)

Variable	Mean	SD	Min	P10	P25	Median	p75	P90	Max
TQ_t	2.359	1.278	0.491	1.225	1.510	1.970	2.775	4.064	7.699
ROE_t	0.120	0.618	−4.471	−0.087	0.082	0.158	0.254	0.411	1.777
$Factor\,1_t$	0.061	0.138	0.000	0.000	0.000	0.000	0.000	0.250	0.750
$Factor\,2_t$	0.072	0.189	0.000	0.000	0.000	0.000	0.000	0.333	1.000
$Factor\,3_t$	0.043	0.162	0.000	0.000	0.000	0.000	0.000	0.000	1.000
$Factor\,4_t$	0.049	0.154	0.000	0.000	0.000	0.000	0.000	0.333	1.000
$Factor\,5_t$	0.044	0.134	0.000	0.000	0.000	0.000	0.000	0.333	1.000
$Factor\,6_t$	0.059	0.126	0.000	0.000	0.000	0.000	0.000	0.250	0.750
$SIZE_{t-1}$	23.528	1.075	21.478	22.221	22.723	23.448	24.222	24.954	26.301
LEV_{t-1}	0.632	0.197	0.137	0.394	0.493	0.632	0.742	0.878	1.232
GRW_t	0.033	0.128	−0.375	−0.106	−0.027	0.025	0.079	0.184	0.645
PPE_{t-1}	0.217	0.301	−0.190	−0.010	0.017	0.063	0.355	0.762	0.917

TQ_t = the sum of the market value of equity and the book value of debt, all divided by the book value of total assets

ROE_t = net income of the current term, all divided by the sum of equity

$Factor\,1_t$ = the average of TF-IDF of the keywords group related to Note to customers

$Factor\,2_t$ = the average of TF-IDF of the keywords group related to Purchasing experience

$Factor\,3_t$ = the average of TF-IDF of the keywords group related to Corporate identity

$Factor\,4_t$ = the average of TF-IDF of the keywords group related to Network infrastructure

$Factor\,5_t$ = the average of TF-IDF of the keywords group related to Supply chain

$Factor\,6_t$ = the average of TF-IDF of the keywords group related to Misc. information

$SIZE_{t-1}$ = natural logarithm result of total assets

LEV_{t-1} = debt ratio (= total debt/total capital)

GRW_t = (sales of this year/sales of last year)−1

PPE_{t-1} = property, plant, and equipment.

<Tables 3>~<Table> 5 and <Table 6> show the results of the regression analysis including the dependent variables for *ROE* and *Tobin's Q*, respectively. In <Table 3>, when the independent variable is a continuum variable, the coefficient value for Factor 1 is −0.498, which is significant and negative at the 5% level. In <Table 4>, when the independent variable is a dummy variable, the coefficient of Factor 1 is −0.155, which is significant and negative at the 10% level. This result indicates that firms with high numbers of words related to *Note to customers* tend to have lower financial performance.

As shown in <Table 5>, when the independent variable is a continuum variable, the

〈Table 3〉 Results for Present Firm Performance

Dependent variable: *ROE*

Independent Variables	Factor = Continuum Variable Full Model			Factor = Continuum Variable Partial Model		
	β	t-value	Significance Level	β	t-value	Significance Level
$Intercept$	−0.858	−0.987	***	−0.708	−0.823	
$Factor\,1_t$	−0.498	−1.860	**	−	−	−
$Factor\,2_t$	0.428	1.891	*	−	−	−
$Factor\,3_t$	−0.219	−0.986		−	−	−
$Factor\,4_t$	0.181	0.690		−	−	−
$Factor\,5_t$	0.116	0.422		−	−	−
$Factor\,6_t$	0.228	0.777		−	−	−
$SIZE_{t-1}$	0.066	1.814	***	0.063	1.756	*
LEV_{t-1}	−0.544	−2.795		−0.603	−3.114	***
GRW_t	0.109	0.339	***	0.033	0.101	
PPE_{t-1}	−0.402	−1.807		−0.463	−2.09	**
$Industry$	Included			Included		
$F-value$	1.512**			1.483*		
Adj_R^2	5.126			3.976		
N_obs	304			304		

ROE_t = net income of the current term, all divided by the sum of equity
$Factor\,1_t$ = the average of TF-IDF of the keywords group related to Note to customers
$Factor\,2_t$ = the average of TF-IDF of the keywords group related to Purchasing experience
$Factor\,3_t$ = the average of TF-IDF of the keywords group related to Corporate identity
$Factor\,4_t$ = the average of TF-IDF of the keywords group related to Network infrastructure
$Factor\,5_t$ = the average of TF-IDF of the keywords group related to Supply chain
$Factor\,6_t$ = the average of TF-IDF of the keywords group related to Misc. information
$SIZE_{t-1}$ = natural logarithm result of total assets
LEV_{t-1} = debt ratio (= total debt/total capital)
GRW_t = (sales of this year/sales of last year)−1
PPE_t−1 = property, plant, and equipment
*, **, *** indicate significance at the 10%, 5%, and 1% levels, respectively.

coefficient of Factor 1 is 1.014, which is positive and significant at the 5% level. In <Table 5>, when the independent variable is a dummy variable, the coefficient of Factor 1 is 0.479, which is positive and significant at the 1% level. This result indicates that firms with a high number of words related to *Note to customers* tend to have relatively higher future corporate value. That is, we can infer that companies with a high number of words related to *Note to customers* in the introductory text tend to be investing in future corporate value rather than current performance.

In <Table 3>, when the independent variable is a continuum variable, the coefficient of Factor 2 is 0.428, which is statistically significant and positive at the 10% level. In <Table 4>, when the independent variable is a dummy variable, the coefficient of Factor 2 is 0.243, which is significant and positive at the 5% level. This result shows the tendency that the higher the number of keywords related to *Purchasing experience* in a firm's institutional information, the higher the firm's financial performance. As shown, in <Table 5>, when the independent variable is a continuum variable, the coefficient of Factor 2 is 0.532, which is not statistically significant. In <Table 5>, when the independent variable is a dummy variable, the coefficient of Factor 2 is 0.213, which is also not significant. This indicates that there is no association between the frequency of keywords related to *Purchasing experience* in the introductory text and future corporate value. It

may therefore be that policies that emphasize purchasing experience may not have a positive effect on future long-term corporate value, although they may affect current performance. The coefficients of the remaining factors, Factors 3, 4, 5, and 6, show no significant relationships with ROE and TQ.

The results for the control variables are as follows. First, firm size (*SIZE*) is positively related to *ROE*, as shown in <Table 3> and <Table 4>, and negatively associated with *Tobin's Q*, as shown in <Table 5> and <Table 6>. These results are consistent with those of Francis et al. [12]: The larger the size of a company's assets, the more likely it is to grow its business based on its wealth of financial resources, which may have a positive impact on corporate performance or future firm value. At the same time, costs may also increase, which may negatively affect the performance or value of the company. For this reason, it seems that the mixed results are found in the analysis.

Second, in <Table 4>, *LEV*, which represents the debt ratio, has a significantly negative relationship with *ROE*. High debt ratios increase the risk of default and at the same time generate high overhead costs. Therefore, the results of this analysis indicate an association between this variable and a decrease in current profit.

In addition, <Table 3>~<Table 6> shows that *GRW*, the variable representing sales growth, has a positive relationship with current

corporate performance and future corporate value. This finding is supported by those of Maitland et al. [28], who argue that higher sales growth has a positive effect on current corporate performance and future corporate value.

The final control variable, PPE, refers to the ratio of tangible assets. In <Table 4>, ROE is converted to a dummy variable with a sig-

nificant negative value. In <Table 6>, *Tobin's Q* is converted to a dummy variable to be comparable with other variables. As the proportion of tangible assets increases, depreciation costs increase [43]. This may have a negative impact on current corporate performance or future corporate value.

In order to test the significance of the six

〈Table 4〉 Results for Present Firm Performance

Dependent variable: ROE

Independent Variables	Factor = Dummy Variable Full Model			Factor = Dummy Variable Partial Model		
	β	t-value	Significance Level	β	t-value	Significance Level
Intercept	−0.761	−0.877		−0.708	−0.823	
Factor 1_t	−0.155	−1.661	*	−	−	−
Factor 2_t	0.243	2.104	**	−	−	−
Factor 3_t	−0.201	−1.451		−	−	−
Factor 4_t	0.158	1.309		−	−	−
Factor 5_t	0.053	0.465		−	−	−
Factor 6_t	0.047	0.515		−	−	−
$SIZE_{t-1}$	0.062	1.693	*	0.063	1.756	*
LEV_{t-1}	−0.559	−2.887	***	−0.603	−3.114	***
GRW_t	0.1	0.311		0.033	0.101	
PPE_{t-1}	−0.404	−1.822*	*	−0.463	−2.09	**
Industry	Included			Included		
F−value	1.555**			1.483*		
Adj_R^2	5.539			3.976		
N_obs	304			304		

ROE_t = net income of the current term, all divided by the sum of equity
Factor 1_t = the average of TF−IDF of the keywords group related to Note to customers
Factor 2_t = the average of TF−IDF of the keywords group related to Purchasing experience
Factor 3_t = the average of TF−IDF of the keywords group related to Corporate identity
Factor 4_t = the average of TF−IDF of the keywords group related to Network infrastructure
Factor 5_t = the average of TF−IDF of the keywords group related to Supply chain
Factor 6_t = the average of TF−IDF of the keywords group related to Misc. information
$SIZE_{t-1}$ = natural logarithm result of total assets
LEV_{t-1} = debt ratio (= total debt/total capital)
GRW_t = (sales of this year/sales of last year)−1
PPE_{t-1} = property, plant, and equipment
*, **, *** indicate significance at the 10%, 5%, and 1% levels, respectively.

factors found in the factor analysis, we conducted multiple regression analysis with reduced model which excluded the six factors. The results are shown in the right hand side of <Table 3>~<Table 6>. When dependent variable is ROE, in <Table 3>, the adjusted R square of the full model is greater than that of the reduced model (full model 5.126%, reduced model 3.976%) when the independent variables are continuum variables. The superiority of the full model is also found in case of dummy variable (see <Table 4>); adjusted R-squared value of the full model and reduced model is 5.539% and 3.976%, respectively.

〈Table 5〉 Results for Future Firm Value

Dependent Variable: *TQ*

Independent Variables	Factor = Continuum Variablev Full Model			Factor = Continuum Variable Partial Model		
	β	t-value	Significance Level	β	t-value	Significance Level
Intercep$_t$	11.485	7.577	***	11.284	7.532	***
Factor 1$_t$	1.014	2.175	**	–	–	–
Factor 2$_t$	0.532	1.351		–	–	–
Factor 3$_t$	0.018	0.046		–	–	–
Factor 4$_t$	0.543	1.188		–	–	–
Factor 5$_t$	−0.473	−0.992		–	–	–
Factor 6$_t$	0.108	0.212		–	–	–
SIZE$_{t-1}$	−0.406	−6.387	***	−0.392	−6.239	***
LEV$_{t-1}$	0.201	0.592		0.167	0.496	
GRW$_t$	2.137	3.813	***	2.14	3.828	***
PPE$_{t-1}$	−0.621	−1.601		−0.627	−1.624	
Industry	Included			Included		
F-value	5.582***			6.451***		
Adj_R^2	32.609			31.870		
N_obs	304			304		

TQ$_t$: = the sum of the market value of equity and the book value of debt, all divided by the book value of total assets
Factor 1$_t$ = the average of TF-IDF of the keywords group related to Note to customers
Factor 2$_t$ = the average of TF-IDF of the keywords group related to Purchasing experience
Factor 3$_t$ = the average of TF-IDF of the keywords group related to Corporate identity
Factor 4$_t$ = the average of TF-IDF of the keywords group related to Network infrastructure
Factor 5$_t$ = the average of TF-IDF of the keywords group related to Supply chain
Factor 6$_t$ = the average of TF-IDF of the keywords group related to Misc. information
SIZE$_{t-1}$ = natural logarithm result of total assets
LEV$_{t-1}$ = debt ratio (= total debt/total capital)
GRW$_t$ = (sales of this year/sales of last year)−1
PPE$_{t-1}$ = property, plant, and equipment;
*, **, *** indicate significance at the 10%, 5%, and 1% levels, respectively.

When it comes to Tobin's Q, as shown in <Table 5>, full model's adjusted R-square is 32.609 % (in case of continuum variable). In the <Table 6>, adjusted R-square of the full model is 34.113 % (in case of dummy variable). The values are greater than reduced models. Consequently, we could conclude that the six factors, especially factor 1 and 2 for ROE, and factor 1 for Tobin's Q, increase the representation power of the regression models. In addition, employing control variables has an effect to show more abundant results.

〈Table 6〉 Results for Future Firm Value

Dependent Variable: *TQ*

Independent Variables	Factor = Dummy Variable Full Model			Factor = Dummy Variable Partial Model		
	β	t-value	Significance Level	β	t-value	Significance Level
$Intercep_t$	11.845	7.906	***	11.284	7.532	***
$Factor\,1_t$	0.479	2.971	***	–	–	–
$Factor\,2_t$	0.213	1.068		–	–	–
$Factor\,3_t$	0.105	0.438		–	–	–
$Factor\,4_t$	0.311	1.497		–	–	–
$Factor\,5_t$	−0.297	−1.517		–	–	–
$Factor\,6_t$	−0.021	−0.129		–	–	–
$SIZE_{t-1}$	−0.419	−6.661	***	−0.392	−6.239	***
LEV_{t-1}	0.195	0.583		0.167	0.496	
GRW_t	2.097	3.778	***	2.14	3.828	***
PPE_{t-1}	−0.677	−1.768	*	−0.627	−1.624	
Industry	Included			Included		
F-value	5.902***			6.451***		
Adj_R^2	34.113			31.870		
N_obs	304			304		

TQ_t: = the sum of the market value of equity and the book value of debt, all divided by the book value of total assets
$Factor\,1_t$ = the average of TF-IDF of the keywords group related to Note to customers
$Factor\,2_t$ = the average of TF-IDF of the keywords group related to Purchasing experience
$Factor\,3_t$ = the average of TF-IDF of the keywords group related to Corporate identity
$Factor\,4_t$ = the average of TF-IDF of the keywords group related to Network infrastructure
$Factor\,5_t$ = the average of TF-IDF of the keywords group related to Supply chain
$Factor\,6_t$ = the average of TF-IDF of the keywords group related to Misc. information
$SIZE_{t-1}$ = natural logarithm result of total assets
LEV_{t-1} = debt ratio (= total debt/total capital)
GRW_t = (sales of this year/sales of last year)−1
PPE_{t-1} = property, plant, and equipment
*, **, *** indicate significance at the 10%, 5%, and 1% levels, respectively.

5. Discussion

5.1 Implications

The results of our empirical analysis have several significant implications. First, keywords related to *Purchasing experience* have a significant and positive relationship with *ROE*, a proxy variable representing current firm performance. This result implies that there is a tendency toward high performance in companies whose website introductory text contains keywords related to purchasing experience. In other words, firms that prioritize and consider customers' purchasing experience in their corporate strategies (as indicated by its importance in the institutional information provided on their websites) perform better than others. For some firms, financial outcomes may improve if the company's website provides information about and shows consideration for customers' purchasing experience. This feeling of consideration is attractive to customers, who tell others, thus fostering loyalty and long-term growth. Helm [17] described the phenomenon of viral marketing, mentioning that a website is an important means of furthering a corporate marketing strategy. Our results are also supported by Stauss [42]'s study, which identified corporate websites as important tools for successful relationship marketing and effective communication.

Second, keywords related to *Note to customers* had a negative relationship with ROE.

This means that firms that emphasize information for consumers in the introductory text on their websites have relatively lower performance. On the other hand, keywords related to *Note to customers* were positively associated with *Tobin's Q*. This result is consistent with that of Hejazi et al. [16] that customer relations is positively associated with Tobin's Q, which is frequently considered as a broad customer-related performance indicator in that in that Tobin's Q is a logical outcome of how well firms are doing with their customers [37]. Well-designed website content, such as *Note to customers*, helps with company-customer communication and eventually results in improved corporate performance.

Cappel and Huang [3] found an association between improved website design, reference links on the homepage, and channels for communicating with customers and better usability. These results deepen our understanding of the relationship between providing customer information, communicating with customers, and firm performance over time. We suggest that it also takes time to see significant financial benefits from providing institutional information on corporate websites. The results of the study of Leporini and Paternò [26] suggest that better performance can ultimately be facilitated by improving website usability in accordance with the firm's vision for its future, and that improvement gradually becomes evident over a long period. In many cases, current finan-

cial performance will suffer due to current expenses, including the cost of building and maintaining a site. However, a company's website is an important marketing tool that can attract new customers and improve customer service [1]. Although some efforts may not improve current performance, they may have a positive effect on future corporate value. The results of our empirical analysis imply that companies emphasizing the keywords related to our Factor 1, *Note to customers*, prioritize communicating actively with customers and providing information. Doing so consistently may increase long-term corporate value that will eventually offset weak current performance.

The academic contribution of this study is as follows. First, this study proves that it is possible to infer firms' financial performance through the institutional information provided in the introductory text on corporate websites. In particular, keywords found in the institutional information on portions of the website such as the "About Us" section can be potentially meaningful indicators of current financial performance and/or future corporate value. Our results may facilitate further exploration of potential determinants of firm performance.

In this study, we suggest an efficient and cost-effective method to obtain information about the present and future performance of a firm using non-financial data. Other methods are time-consuming and challenging. For example, data for Tobin's Q, a typical corporate value indicator, are often not available or hard to obtain. For accurate calculation of this indicator, a complex process is necessary using a variety of financial data. This remains as further research issue.

Third, our results emphasize the importance of institutional information revealed in corporate websites. The results of the empirical analysis imply that companies should be more careful when they convey information through their websites. At the same time, users can obtain useful information from a firm's website, particularly regarding its current and future performance.

Last, the results of our study indicate the usefulness of the text mining method based on big data analytics in finance and accounting research. Collecting data for empirical studies in these fields can be difficult. We showed that the text mining method can formalize unstructured non-financial data like the institutional information provided on websites. This new method converges analytical methods from other fields, enabling work that was not previously possible.

5.2 Conclusion

Most companies build and maintain websites that contain a variety of information [25]. The introductory message and institutional information on corporate websites are especially useful sources from which interested parties may gain insight into the current and future performance of a firm. Our analysis of the rela-

tionship between the introductory text provided in the "About Us" section and current firm performance and future corporate value involved evaluation of keywords extracted from this institutional information. The results of this paper provide insights that cannot be obtained from quantitative data which are acquired from open firm data set such as Osiris. Moreover, this study reconfirms the usefulness of text analysis in e-commerce research area [7, 20].

The limitations of this paper are as follows. First, we use data from a single year due to the difficulty of gathering data. Second, this study's sample includes U.S. S&P500 companies. Expansion of the study period and sample would facilitate generalization of the results. In addition, future studies may include other pieces of information on corporate websites besides the introductory text.

References

[1] Acharya, R. N., Kagan, A., Lingam, S. R., and Gray, K., "Impact of website usability on performance: A heuristic evaluation of community bank homepage implementation," Journal of Business and Economics Research, Vol. 6, No. 6, pp. 139-148. 2008.

[2] Akindele, J. and Odusina, O., "Working capital management and firm profitability: Evidence from Nigeria quoted companies," Journal of Finance and Accounting, Vol. 6, No. 7, pp. 148-153, 2015.

[3] Cappel, J. J. and Huang, Z., "A usability analysis of company websites," Journal of Computer Information Systems, Vol. 48, No. 1, pp. 117-123, 2007.

[4] Chadwick, C., Super, J. F., and Kwon, K., "Resource orchestration in practice: CEO emphasis on SHRM, commitment-based HR systems, and firm performance," Strategic Management Journal, Vol. 36, No. 3, pp. 360-376, 2015.

[5] Chen, K., Zhang, Z., Long, J., and Zhang, H., "Turning from TF-IDF to TF-IGM for term weighting in text classification," Expert Systems with Applications, Vol. 66, pp. 245-260, 2016.

[6] Choi, S. and Kwon, O., "The study of developing Korean SentiWordNet for big data analytics: Focusing on anger emotion," The Journal of Society for e-Business Studies, Vol. 19, No. 4, pp. 1-19, 2014.

[7] Choi, S., Lee, J., and Kwon, O., "A morphological analysis method of predicting place-event performance by online news titles," The Journal of Society for e-Business Studies, Vol. 21, No. 1, pp. 15-32, 2016.

[8] Chung, K. H. and Pruitt, S. W., "A simple approximation of Tobin's q," Financial Management, Vol. 23, No. 3, pp. 70-74, 1994.

[9] Cornelissen, J. and Cornelissen, J. P., Corporate Communication: A Guide to

Theory and Practice, Sage, 2017.

[10] Fadda, N. and Sørensen, J. F. L., "The importance of destination attractiveness and entrepreneurial orientation in explaining firm performance in the Sardinian accommodation sector," International Journal of Contemporary Hospitality Management, Vol. 29, No. 6, pp. 1684-1702, 2017.

[11] Fauver, L., Hung, M., Li, X., and Taboada, A. G., "Board reforms and firm value: Worldwide evidence," Journal of Financial Economics, Vol. 125, No. 1, pp. 120-142, 2017.

[12] Francis, J. R., Pinnuck, M. L., and Watanabe, O., "Auditor style and financial statement comparability," The Accounting Review, Vol. 89, No. 2, pp. 605-633, 2014.

[13] Geissler, G. L., Zinkhan, G. M., and Watson, R. T., "Web home page complexity and communication effectiveness," Journal of Association for Information Systems, Vol. 2, pp. 1-46, 2001.

[14] Glaveli, N. and Geormas, K., "Doing well and doing good: Exploring how strategic and market orientation impacts social enterprise performance," International Journal of Entrepreneurial Behavior & Research, Vol. 24, No. 1, pp. 147-170, 2018.

[15] Hasan, B., "Perceived irritation in online shopping: The impact of website design characteristics," Computers in Human Behavior, Vol. 54, pp. 224-230, 2016.

[16] Hejazi, R., Ghanbari, M., and Alipour, M., "Intellectual, human and structural capital effects on firm performance as measured by Tobin's Q," Knowledge and Process Management, Vol. 23, No. 4, pp. 259-273, 2016.

[17] Helm, S., "Viral marketing-establishing customer relationships by 'word-of-mouse'," Electronic Markets, Vol. 10, No 3, pp. 158-161, 2000.

[18] Hwang, J. S., McMillan, S. J., and Lee, G., "Corporate web sites as advertising: An analysis of function, audience, and message strategy," Journal of Interactive Advertising, Vol. 3, No. 2, pp. 10-23, 2003.

[19] Kim, J. and Kim, D., "A study on the method for extracting the purpose-specific customized information from online product reviews based on text mining," The Journal of Society for e-Business Studies, Vol. 21, No. 2, pp. 151-161, 2016.

[20] Kim, J. Y. and Kim, D., "A Study on the method for extracting the purpose-specific customized information from online product reviews based on text mining," The Journal of Society for e-Business Studies, Vol. 21, No. 2, pp. 151-161, 2017.

[21] Kim, W. G., Lim, H., and Brymer, R. A., "The effectiveness of managing social media on hotel performance," International Journal of Hospitality Management, Vol. 44, pp. 165-171, 2015.

[22] Kluge, P. N., Königsfeld, J. A., Fassnacht, M., and Mitschke, F.. "Luxury web atmospherics: an examination of homepage de-

sign," International Journal of Retail & Distribution Management, Vol. 41, No. 11/12, pp. 901–916, 2013.

[23] Lee, D., Kim, Y., and Kim, K., "Topic based hierarchical network analysis for entrepreneur using text mining," The Journal of Society for e-Business Studies, Vol. 23, No. 3, pp. 33–49, 2018.

[24] Lee, K. H., Cin, B. C., and Lee, E. Y., "Environmental responsibility and firm performance: the application of an environmental, social and governance model," Business Strategy and the Environment, Vol. 25, No. 1, pp. 40–53, 2016.

[25] Lee, S. K., Son, J. S., and Kim, K., "A method for calculating exposure risks of privacy information based on website structures," Journal of Society for e-Business Studies, Vol. 21, No. 1, pp. 1–14, 2016.

[26] Leporini, B. and Paternò, F., "Applying web usability criteria for vision-impaired users: does it really improve task performance," International Journal of Human-Computer Interaction, Vol. 24, No. 1, pp. 17–47, 2008.

[27] Li, L., "The effects of trust and shared vision on inward knowledge transfer in subsidiaries' intra-and inter-organizational relationships," International Business Review, Vol. 14, No. 1, pp. 77–95, 2005.

[28] Maitland, E., Rose, E. L., and Nicholas, S., "How firms grow: clustering as a dynamic model of internationalization," Journal of International Business Studies,

Vol. 36, No. 4, pp. 435–451, 2005.

[29] Martini, A., Massa, S., and Testa, S., "The inextricable intertwining of the firm, the platform and the customer: the case of a social media platform for innovation," Empowering Organizations, Vol. 11, pp. 115–131, 2016.

[30] Mishra, K. E., "Help or hype: Symbolic or behavioral communication during Hurricane Katrina," Public Relations Review, Vol. 32, No. 4, pp. 358–366, 2006.

[31] Moon, N., Opinion Polls: History, Theory and Practice 1st edition Manchester: Manchester University Press, 1999.

[32] Nguyen, N. P., Ngo, L. V., Bucic, T., and Phong, N. D., "Cross-functional knowledge sharing, coordination and firm performance: The role of cross-functional competition," Industrial Marketing Management, Vol. 71, pp. 123–134, 2017.

[33] Park, S. O., Na, H. J., and Kwon, O., "Comparative effect of company-driven SNS activity vs. consumer-driven SNS activity on firm value: Evidence from Facebook," Computers in Industry, Vol. 82, pp. 186–195, 2016.

[34] Porter, M. F., "An Algorithm for Suffix Stripping," Program, Vol. 14, No. 3, pp. 130–137, 1980.

[35] Rosenbaum, H., "The information environment of electronic commerce: information imperatives for the firm," Journal of Information Science, Vol. 26, No. 3, pp. 161–171, 2000.

[36] Ross, S., "The determinants of financial structure: The incentive signalling approach," Bell Journal of Economics, Vol. 8, pp. 23-40, 1977.

[37] Rubera, G. and Kirca, A. H., "You gotta serve somebody: the effects of firm innovation on customer satisfaction and firm value," Journal of the Academy of Marketing Science, Vol. 45, No. 5, pp. 741-761, 2017.

[38] Ruzza, M., Tiozzo, B., Mantovani, C., D'Este, F., and Ravarotto, L., "Designing the information architecture of a complex website: A strategy based on news content and faceted classification," International Journal of Information Management, Vol. 37, No. 3, pp. 166-176, 2017.

[39] Sengur, Y. and Sengur, F. K., "Airlines define their business models: a content analysis," World Review of Intermodal Transportation Research, Vol. 6, No. 2, pp. 141-154, 2017.

[40] Servaes, H. and Tamayo, A., "The impact of corporate social responsibility on firm value: The role of customer awareness," Management Science, Vol. 59, No. 5, pp. 1045-1061, 2013.

[41] Srinivasan, S. and Hanssens, D. M., "Marketing and firm value: Metrics, methods, findings, and future directions," Journal of Marketing Research, Vol. 46, No. 3, pp. 293-312, 2009.

[42] Stauss, B., Using new media for customer interaction: A challenge for relationship marketing, Relationship Marketing, Springer Berlin Heidelberg, 2000.

[43] Villalonga, B. and Amit, R., "How do family ownership, control and management affect firm value," Journal of Financial Economics, Vol. 80, No. 2, pp. 385-417, 2006.

[44] Wang, D. H. M., Chen, P. H., Yu, T. H. K., and Hsiao, C. Y., "The effects of corporate social responsibility on brand equity and firm performance," Journal of Business Research, Vol. 68, No. 11, pp. 2232-2236, 2015.

저 자 소 개

나형종 (E-mail: fresh_na_77@hanmail.net)
2016년 경희대학교 경영학(회계학) (박사)
2017년 경희대학교 경영대학 연구교수
2018년~현재 성균관대학교 경영대학 연구교수
관심분야 Text mining과 Aspect mining을 통한 기업의 재무·조세 분석 및 예측

최석재 (E-mail: sjchoi@khu.ac.kr)
2001년~2002년 카네기멜론대학 전산학부 방문연구원
2003년~2005년 중국연변과기대 언어공학연구소 실장
2008년 고려대학교 국어국문학과 (박사)
2014년~현재 경희대학교 빅데이터연구소 및 BK21 학술연구교수
관심분야 비정형데이터, 빅데이터분석, 챗봇시스템

권오병 (E-mail: obkwon@khu.ac.kr)
1995년 한국과학기술원 경영과학과 (박사)
1996년~2004년 한동대학교 경영경제학부 부교수
2004년~현재 경희대학교 경영대학 교수
관심분야 빅데이터, 인공지능 경영, IoT