

토픽모델링 기법을 활용한 산업별 직무만족요인 비교 조사 : 잡플래닛 리뷰를 중심으로*

김동욱** · 강주영*** · 임재익****

Comparative Analysis of Job Satisfaction Factors, Using LDA Topic Modeling by Industries : The Case Study of Job Planet Reviews*

Dongwook Kim** · Juyoung Kang*** · Jay Ick Lim****

■ Abstract ■

As unemployment rates and concerns about turnover keep growing, the need for information is also increasing. In these situations, the job reviews which share information about the company catch people's attention because they are usually created by people who worked at the company. The development of SNS and mobile environments has led to an increase in the web services that provide job reviews. For example, Jobplanet is a job review service in Korea, and Glassdoor.com offers a similar service in the US. Despite this attention, however, research utilizing job reviews is insufficient. This paper asks whether there are differences in ratios of job satisfaction factors by industry, using LDA topic modeling and co-occurrence analysis to explore the differences. Through the results of LDA, we find that the ratios of job satisfaction factors are similar by industry. At the same time, the results of co-occurrence analysis show that the co-occurrence frequency of some job satisfaction factors appears high: pay and welfare, balance of work and life, company culture. We expect that the result of this research will be helpful in comparative analysis of job satisfaction factors by industry. Furthermore, in this paper we suggest how to use the job review data in organizational behavior research.

Keyword : LDA Topic Modeling, Facet job satisfaction, Job Review, Co-occurrence Analysis

Submitted : April 29, 2016

1st Revision : July 22, 2016

Accepted : July 25, 2016

* 본 과제(결과물)는 교육부의 재원으로 지원을 받아 수행된 산학협력 선도대학(LINC) 육성사업의 연구결과입니다.

** 아주대학교 e-비즈니스학과, 학사과정

*** 아주대학교 e-비즈니스학과 교수, 교신저자

**** 아주대학교 e-비즈니스학과 부교수

1. 서 론

최근 SNS와 스마트폰의 발달로 인터넷 사용자들이 직접 제작하여 온라인에 올리는 UGC(User-generated content)에 대한 관심이 점점 커지고 있다. UGC는 초기에는 글과 사진 위주의 엔터테인먼트 콘텐츠 형태에서 동영상 위주의 정보 제공 콘텐츠 위주로 발전하고 있다(Doopedia, 2016). 따라서, 많은 인터넷 사용자들은 UGC를 통해 다양한 정보를 얻고 있고, 기업들은 이를 마케팅에 활용하기 위한 방안들을 고민하고 있다. 특히 기업 정보에 관한 UGC가 많은 관심을 받고 있는데, 전 세계적으로 장기적인 취업난이 심화되고 고용 불안정으로 인해 이직에 대한 관심이 높아짐에 따라 다양한 기업 정보에 대한 관심이 증가하고 있다.

최근 영국에서 이루어진 Lakin(2015)의 연구에 따르면 사람들이 기업에 대한 정보를 얻을 때 기업이나 국가에서 제공하는 직업 가이드라인과 같은 전통적인 정보보다 실제 그 기업에서 근무한 경험이 있는 직원들이 작성한 기업 리뷰를 제공하는 글래스도어(Glassdoor.com)와 같은 웹 서비스를 더 신뢰한다고 밝혀졌다. 또한, 68%에 달하는 응답자들이 가족, 친구, 직장동료로부터 구전되는 정보(Word of Mouth) 다음으로 글래스도어와 같은 웹 서비스를 통하여 얻는 정보를 신뢰한다고 응답하였다. 글래스도어는 미국의 기업 정보에 관한 전/현직자들의 리뷰를 익명으로 제공하는 웹 서비스로서, 기업 리뷰뿐만 아니라 연봉 정보, 면접 리뷰 등 구직자에게 필요한 다양한 정보들을 제공하고 있다. 한국의 경우 글래스도어를 벤치마킹한 잡플래닛(Jobplanet.co.kr)이란 웹 서비스가 있는데, 잡플래닛 또한 글래스도어와 같이 기업 리뷰, 연봉 정보, 면접 리뷰 등을 제공하고 있다.

Lakin(2015)의 연구에 따르면 글래스도어나 잡플래닛과 같은 웹 서비스가 최근 고용시장에서 상당히 중요한 역할을 수행한다는 점을 보여주고 있다. 많은 사람들이 글래스도어와 같은 웹 서비스

를 신뢰한다는 사실은 이 플랫폼이 구직자가 기업을 결정할 때 많은 영향을 주고 있음을 나타내고 있다.

그렇다면 왜 많은 사람들은 글래스도어 또는 잡플래닛과 같은 기업 리뷰 서비스에 열광하는 것일까? 전통적인 기업 정보와 달리 온라인상에서 제공하는 기업 리뷰는 익명을 기반으로 실제 근무 경험이 있는 사람의 기업에 대한 솔직한 생각을 보여주고 있다. 이러한 기업 리뷰의 특징은 직무만족연구에 있어서 기업 리뷰를 이용하면 직원들의 심층적인 심리를 파악할 수 있는 가능성을 보여준다. 하지만 기업 리뷰 데이터를 활용한 실증적인 연구들은 아직까지 활발히 이루어지지 않고 있다. 국외의 경우 글래스도어의 기업 리뷰 데이터를 이용하여 허츠버그의 동기-위생이론에 존재하는 주요한 요인들에 대한 타당성 검증을 한 연구가 있다(Dekay, 2013). 이에 비해 국내는 기업 리뷰 데이터를 이용한 연구는 아직까지 찾아보기 어려운 실정이다.

이에 본 연구는 한국의 글래스도어인 잡플래닛의 리뷰 데이터에 대해 텍스트 분석 방법론인 LDA 토픽모델링 기법과 동시출현 단어분석 기법을 도입하여 산업별로 직무 만족 요인에 차이가 있는 지 탐색적으로 살펴보고자 한다. 본 연구를 통해 다음과 같은 사실을 기대할 수 있을 것이다. 첫째로 데이터 수집의 어려움으로 인해 실시하지 못 했던 산업 간 직무만족요인의 비교를 할 수 있을 것이다. 둘째로 본 연구는 직무만족 연구뿐만 아니라 조직행동연구에 있어서 리뷰 데이터를 이용한 사례가 될 수 있을 것으로 기대된다.

본 연구의 구성은 다음과 같다. 우선 제 2장에서는 리뷰 데이터 분석에 이용되는 LDA 토픽모델링과 동시출현 단어분석에 대해 설명하고 직무만족이론에 대한 선행연구에 대하여 서술한다. 제 3장에서는 구체적인 연구 절차 및 방법에 대하여 서술한다. 제 4장은 연구에 대한 분석 및 결과를 살펴보고, 마지막으로 제 5장에서는 본 연구의 시사점 및 한계에 대하여 토의하고자 한다.

2. 선행연구

2.1 LDA 토픽모델링

LDA(Latent Dirichlet Allocation) 토픽모델링은 다수의 비구조적인 문서(Unstructured document)에서 글의 주제와 관련된 정보들을 기반으로 문서를 분류하는 확률 모형이다(Blei et al., 2003). LDA 방법론은 문서가 작성되는 방식을 다음과 같이 가정하고 있다. 먼저 어떤 분포를 따르는 k 개의 토픽이 존재한다고 가정할 때 이들 중 임의로 특정한 토픽을 선택되며, 각각의 토픽들은 다양한 단어가 출현할 확률을 갖고 있다. 이 때, 문서는 단어가 출현할 확률 분포에 맞추어서 임의로 특정한 단어를 선택하는 방식으로 구성된다(Blei, 2012).

위와 같은 방식으로 문서가 구성될 때 우리는 결과 문서만 알고 있고, 토픽에 대한 확률 분포나 토픽에 따른 단어들의 확률 분포 등 토픽을 형성하는 구조들에 대해서는 모르고 있다. LDA 토픽모델링은 이러한 상황에서 확률적으로 숨겨진 구조(hidden structure)를 찾을 수 있게 해주는 확률론적인 모델링(probabilistic modeling) 방법이며 다음과 같은 조건부 확률 분포 식(conditional probabilistic distribution)을 갖는다.

$$p(\beta_{1:K}, \theta_{1:D}, z_{1:D}, w_{1:D}) \\ = \prod_{i=1}^K p(\beta_i) \prod_{d=1}^D p(\theta_d) \\ \left(\prod_{n=1}^N p(z_{d,n} | \theta_d) p(w_{d,n} | \beta_{1:K}, z_{d,n}) \right)$$

β_k (k th topic on documents)

θ_d (Topic proportions on d th document)

$z_{d,n}$ (Topic assignment n th words on d th document)

$w_{d,n}$ (Observed n th word on d th document)

<Figure 1> LDA Probability Distribution(Blei, 2012)

<Figure 1>은 관찰된 데이터 $w_{d,n}$ 를 이용해서 모수(parameter) $\beta_k, \theta_d, z_{d,n}$ 을 추정하는 알고리즘이다. LDA 토픽모델링은 예전의 키워드 중심으로 텍스트를 분류하는 방식에 비하여 훨씬 효과적으로 문서들을 분류할 수 있어서 국, 내외 다양한 분야에서 사용되어 왔다. 국외의 경우 Grimmer (2010)는 미 상원의원들이 주장하는 정책들의 우선순위를 효과적으로 파악하고 측정하기 위해 토픽모델링을 활용했다. 또한 Zhao et al.(2011)는 트위터(twitter)가 전통적인 미디어와 비교해서 모든 뉴스에 대해 빠른 정보 전달이 이루어지는지 비교 분석하기 위해 토픽모델링을 이용했다. 국내 연구의 경우 Park and Song(2013)은 문헌 정보학 분야에서 연구동향을 파악하기 위하여 국내 주요 학술지에 등재된 최근 30년의 논문 초록에 대한 토픽모델링 분석을 실시했다. 또한 최근 소셜 데이터(social data)를 활용하여 마케팅 전략을 수립하는 데 토픽모델링 기법이 이용되고 있다(Cha et al., 2015; Lee, 2015). 이 외에도, Chae et al. (2015)는 모바일 앱의 리뷰 데이터를 이용하여 소셜커머스와 오픈마켓의 이용경험을 비교 분석하기 위해 토픽모델링을 활용하였다.

이에 본 연구에서는 기업 리뷰 데이터를 심층적으로 분석하기 위하여 토픽모델링을 기법을 도입하고자 한다. 먼저 기업 리뷰 데이터를 총 30개의 토픽으로 분류하고 이들 토픽들의 특성을 파악하여 직무만족요인으로 재분류 할 것이다. 이를 통해 산업 간 직무만족요인이 어떻게 다른지 파악하고 조직연구에서 기업 리뷰 데이터가 어떻게 활용 가능한지 논의하고자 한다.

2.2 동시출현 단어분석

1980년대 Callon et al.(1983)이 소개한 동시출현 단어분석(Co-occurrence Network)은 네트워크 분석 중 하나의 방법으로써 동시출현 단어들을 이용하여 문서들 간의 주제 유사도를 측정하는 기법이다. 동시출현 단어분석은 동시출현행렬(Co-oc-

currence matrix)를 그래프로 시각화하여 그래프에서 단어 간의 연관성에 대해서 나타낼 수 있다(Callon et al., 1983). 동시출현 단어분석에서 단어의 출현 횟수가 많을수록 그래프의 중앙에 상대적으로 큰 원으로 나타나고 단어의 출현 횟수가 적으면 그래프의 가장자리에 상대적으로 작은 원으로 나타나게 된다. 만약 두 단어의 동시출현횟수가 많을수록 두 노드(node)를 연결하는 링크(link)의 굵기가 굵어지고 반대로 두 단어의 동시출현횟수가 적을수록 두 노드를 연결하는 링크의 굵기가 가늘게 나타난다(Callon et al., 1983). 이와 같은 동시출현 단어분석 기법은 주로 특정 학문의 지적구조, 연구 주제, 연구동향 등을 파악하는데 주로 사용되고 있다(Kim et al., 2015a; Kim and Song, 2014)

특히, 최근에는 토픽모델링과 동시출현 단어분석을 함께 사용하여 토픽의 구조 및 관계를 심층적으로 파악하는 연구들이 등장하고 있다. Jin et al. (2013)는 트위터(twitter)에서 토픽의 변화 시점 및 패턴을 파악하기 위해 동시출현 단어분석을 실시하고 이를 검증하기 위하여 LDA 토픽모델링을 실시했다. 이 밖에도 Lee(2015)는 그래프 및 토픽모델링 기법을 이용하여 댓글 그래프 개념을 제안하고 이를 기반으로 이슈를 트래킹 하는 기법을 소개하였고, Kim et al.(2015b)은 댓글의 특성을 분류하는 기법에 관해, Lee(2015)는 유용한 리뷰를 식별할 수 있는 방안을 연구하였다.

본 연구는 동시출현 단어분석 기법을 이용하여 토픽들로 구성된 직무만족요인들의 관계 및 토픽 내의 단어들의 관계를 살펴보고자 한다.

2.3 직무만족도 및 직무만족요인

직무만족도(Job satisfaction)은 개인이 직무에 대해 갖고 있는 태도로서, 직무와 관련하여 다양한 경험을 평가했을 때 발생하는 만족감 혹은 긍정적인 감정 상태를 의미한다(Locke, 1976). 직무만족도는 1930년대부터 연구하기 시작한 개념으로서, 조직연구분야에서 중요한 개념이지만 구성개념적

특성이 강하기 때문에 가장 이론적 기반이 취약한 개념 중 하나다(Bussing et al., 1999; Cho and Ju, 2001).

이러한 직무만족도를 측정하기 위한 다양한 방법들이 연구되어 오고 있으며, 크게 설문조사를 이용한 방법과 인터뷰를 이용한 방법으로 이루어지고 있다. 설문조사 방법의 경우 주로 우선순위법과 같은 도구를 이용하여 진행한다. 대표적인 우선 순위법은 MSQ(Minnesota Satisfaction Questionnaire), JDI(Job Descriptive Index), NSQ(Need Satisfaction Questionnaire)와 같은 방법들이 있다(Scarpello and Campbell, 1983). 이 방법들은 구체적인 직무 상황, 직무와 관련된 실질적 요소 또는 상황에 대해서 5~7점의 리커트(Likert) 척도나 '예/아니오'와 같은 척도로 측정한다. 설문조사는 질문사항이 간단하여 이해가 용이하고 대담이 쉽다는 장점을 가지고 있다(Cho and Ju, 2001). 반면 설문조사는 질문에 관한 정보들만 수집가능하기 때문에 제한된 정보만 측정할 수 있는 한계가 존재한다(Coughlan et al., 2009).

이에 비해 인터뷰를 이용한 방법은 기존의 질문자가 잘 이해하지 못하는 상황이나 내용에 대해서 알 수 있고 응답자의 보다 심층적인 심리에 대해서 살펴볼 수 있다는 장점이 있다(Cranny et al., 1992). 하지만 이러한 장점에도 불구하고 설문조사에 비해 시간과 비용이 많이 들기 때문에 연구에서 인터뷰를 이용하는데 적지 않은 어려움이 따랐다. 그러나 웹 사이트에서 인터뷰 형식으로 구성된 기업 리뷰 데이터를 활용하면 인터뷰의 한계점을 극복할 수 있을 것으로 기대된다. 또한, 기업 리뷰데이터는 다양한 기업에 대한 정보를 수집할 수 있기 때문에 표본에 대한 대표성을 향상시킬 수 있고, 기업에 대한 근로자의 솔직한 생각을 알 수 있다는 점에서 장점이 있다. 따라서, 본 연구에서는 기존 연구에서 실시한 측정의 한계점을 보완하면서 향상된 결과를 도출하기 위해 기업 리뷰 데이터를 이용한 분석을 실시하고자 한다.

본 연구에서 측정하는 직무만족요인은 많은 연

구에서 활용되고 있는 척도인 JDI(Job Descriptive Index)를 활용하고자 한다(Cho and Ju, 2001). 1960년대에 개발된 JDI는 직무와 관련된 실질적인 요소에 대한 질문을 수행하는 척도로서 많은 선행 연구를 통해 유용성이 검증되어 오고 있다(Jung and Kim, 2008).

본 연구에서는 잡플래닛에서 평점의 기준으로 사용되는 JDI 척도와 많은 선행연구에서 중요하게 판단하는 척도를 고려하여 직무만족요인을 선택하였다. 그 결과 직무내용(Work Itself, WI), 승진기회 및 가능성(Career Opportunities, CO), 급여 및 복지(Compensation and Benefits, CB), 업무와 삶의 균형(Work/Life Balance, WLB), 사내문화(Culture and Values, CV), 경영진(Senior Management, SM)이라는 총 6가지 직무만족요인을 선택하였다. 먼저, 직무내용의 경우 직무의 내용이 개인의 학습, 흥미, 발전과 성장의 기회, 책임감 등을 제공하는가에 관한 요인으로서, 많은 연구에서 직무내용은 직무 만족도와 높은 상관관계를 갖는 것으로 드러났다(Kinicki et al., 2002). 두 번째로 승진기회 및 가능성은 조직 내 이루어지는 승진의 방법에 따라서 직무만족에 미치는 영향에 관한 요인으로, 만약 회사의 승진 방식이 연공서열에 근거한 방식이면 높이는 것으로 알려져 있다. 셋째, 급여 및 복지는 직접적 보상인 금전과 간접 보상인 후생 복지가 얼마나 직무만족도에 영향을 주는지 나타내는 요인이다. 넷째, 업무와 삶의 균형은 직위가 얼마나 안정적인가 일과 개인의 삶이 얼마나 조화를 이루는가를 측정하는 요인이다. 다섯째, 사내 문화는 회사 내의 문화적으로 존재하는 동료와 상사의 영향을 보여주는 요인이다. 마지막으로 경영진은 회사의 정책 또는 경영진들의 선택에 따라서 직무만족도에 영향을 미치는 요인이다(Ricketta, 2008).

2.4 산업별 직무만족요인 연구

지금까지 직무만족요인에 관하여 많은 연구가

있었다. Cho and Ju(2001)는 기업의 공공성에 따라서 직업만족도와 직업만족에 영향을 미치는 요인의 차이에 대하여 분석을 했다. 또한 Park et al. (2009)은 미디어 기업의 창의적 조직문화라는 요인이 다른 직무만족요인들과 비교하여 직무만족도에 얼마나 영향을 미치는 지 연구를 진행했다. 특히 Clark(1996)의 경우 영국인의 직무만족도를 구성하는 다양한 요인들을 대한 연구를 진행함으로써 영국인의 직무만족도를 측정하는 데 중점이 되는 요인을 발견했다.

그러나 다양한 산업의 직무만족도 연구에도 불구하고 산업 간의 직무만족요인에 대한 비교 연구가 이루어지지 않았다. 이는 일관된 기준을 바탕으로 서로 다른 산업 간의 데이터를 측정하는데 많은 어려움이 있기 때문이라고 추측되는데, 만약 설문조사를 통해 측정할 경우 산업별로 특이성이 존재하기 때문에 하나의 설문지를 통해 조사하기 어렵다. 또한 다양한 산업에 걸친 기업에 대하여 조사를 실시해야 되기 때문에 상당한 비용적 문제가 발생한다. 하지만 기업 리뷰 데이터를 이용한다면, 산업 간 리뷰를 쉽게 분류 및 수집을 할 수 있기 때문에 산업 간 직무만족요인의 비교 분석이 가능할 것으로 판단되어 본 연구에서는 산업별 직무 만족에 대해 탐색 연구를 수행하고자 한다.

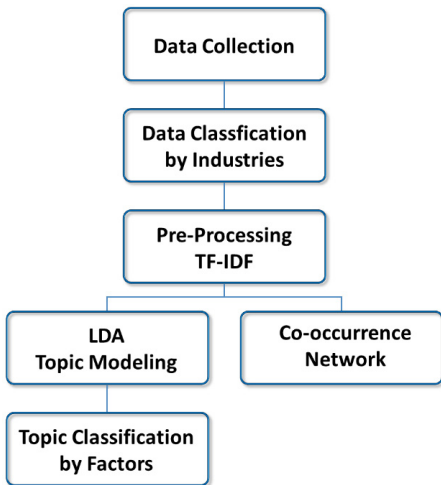
본 연구에서 살펴볼 산업은 잡플래닛에서 분류한 총 10개의 산업을 기준으로 살펴볼 것이다. 각각의 분류는 서비스업(Service, S), 제조/화학(Manufacture/Chemical, MC), 의료/제약(Medical/ Pharmaceutical, MP), 유통/무역/운송(Distribution/Trade/Transportation, DTT), 교육(Education, E), 건설업(Architecture, A), IT(IT, IT), 미디어/디자인(Media/Design, MD), 금융업(Finance, F), 기관/협회(Organization/ Association, OA)이다.

3. 연구 절차 및 방법

본 연구는 산업별 직무 만족 요인을 비교하기 위하여 기업 리뷰 데이터를 심층적으로 비교 분석

하는 탐색적 연구를 수행하고자 한다. 이를 위해, 잡플래닛이라는 기업 리뷰 사이트로부터 데이터를 수집하여, 산업별 직무 만족도에 관해 토픽모델링 및 동시출현 단어 분석 기법을 활용하고자 한다.

본 연구에서 수행되는 절차는 <Figure 2>와 같다. 먼저 잡플래닛의 실제 기업 리뷰데이터를 웹 크롤링 방법을 통해 수집한다. 다음 단계로 수집된 데이터는 잡플래닛에서 구분하는 기준에 따라 10개의 산업을 기준으로 분류된다. 세 번째 단계로, 분류된 리뷰 데이터는 형태소 분석 및 불용어 제거와 같은 텍스트 전처리 과정을 거칠 것이다. 그 후 전 처리된 리뷰 데이터를 LDA 토픽모델링을 통하여 토픽을 추출한 뒤 해당 토픽을 직무만족요인으로 재분류하여 산업별 직무만족요인의 차이를 살펴볼 것이다. 또한 동시출현단어분석을 통해 토픽 내 단어들 사이의 관계에 대해서 심도 있게 살펴볼 것이다.



<Figure 2> Overall Research Process

3.1 기업 리뷰 데이터

본 연구에서는 온라인에서 수집된 기업 리뷰 데이터를 이용하여 연구를 진행하였다. 직무 만족도를 살펴보는 기존 연구들은 대부분 설문조사를 이용한 데이터 수집에 의존하고 있다. 하지만 설문조사를 통한 연구는 다음과 같은 한계점을 가지고

있다. 먼저 설문조사는 사회적 적절성 편견 (Social desirable bias)이 발생하여 응답자의 솔직한 의견과 다른 거짓으로 답변할 가능성이 있다(Demaio, 1984). 예를 들면 근로자의 직무만족도에 대하여 설문조사를 할 경우 근로자는 자신의 응답이 불이익으로 작용할 수 있다는 걱정 때문에 솔직하지 않은 응답을 할 수 있다. 둘째로 설문조사는 미리 측정할 변수들에 대한 질문을 설정하기 때문에 제한된 결과만 수집된다는 한계점이 있다(Coughlan et al., 2009). 이에 비해 기업 리뷰 데이터는 익명성이 보장되어 직원들의 기업에 대한 솔직한 의견을 알 수 있고, 다양한 분석 결과를 얻을 수 있다는 장점이 있다(Chae et al., 2015).

이러한 한계를 극복하고자 본 연구에서는 리뷰 데이터를 이용한 직무만족연구를 실시하고 이를 통해 조직 연구에서 리뷰 데이터의 활용 가능성에 대하여 제시해 보고자 한다.

3.2 텍스트 분석 과정

기업 리뷰 데이터들은 분석을 위해 띄어쓰기, 문법 오류, 사용자의 실수로 인한 오타 등으로 인해 적절한 텍스트 전 처리 과정이 요구된다. 본 연구는 전처리된 리뷰들을 토픽을 기준으로 분류하기 위하여 LDA 토픽모델링을 사용하였다. LDA는 문서를 주제들의 확률적인 합으로써 인식하고 주제들을 기준으로 분류하기 때문에 다른 분류법에 비해 효과적으로 문서를 분류할 수 있다는 장점이 있다(Chae et al., 2015).

본 연구에서는 R의 lda 패키지를 이용하여 각 산업별 리뷰 데이터에 대하여 30개의 토픽으로 분류하고 분류된 토픽은 JDI의 6개 요인들로 재분류하였다. 많은 선행 연구에서 lda를 통해 분류된 토픽들 간의 공통점을 파악하여 주된 토픽을 파악하였다. 예를 들면 Chae et al.(2015)는 트위터에 대하여 lda 토픽모델링으로 10개의 토픽으로 분류한 후 서로 다른 토픽들을 lifestyle이나 expectation이라는 공통적인 속성으로 표현하였다. 또한 Herzberg

(1968)는 직무의 어떤 요인이 만족 혹은 불만족을 일으키는 지 살펴보기 위하여, 인터뷰를 통해 데이터를 수집하고 이를 주제별로 분류하여 직무만족요인을 파악하였다.

이와 같이 기존의 선행 연구들은 다양한 토픽을 통하여 하나의 구성적 개념을 표현할 수 있음을 시사한다. 특히, 잡플래닛의 총평 같은 경우 리뷰어가 회사에서 겪은 직무에 대한 경험에 관하여 구체적으로 표현할 수 있다. 즉 리뷰어가 기업에 대하여 구체적으로 어떤 요인들이 만족 또는 불만족을 일으키는 지 요약적으로 서술할 수 있다. 그러므로 이러한 총평에서 드러나는 토픽을 통해 직무만족요인에 대한 구성적 개념을 표현할 수 있다.

보통 토픽을 탐색적으로 분류하지만 조직만족요인의 경우 매우 다양한 요인들이 존재하기 때문에 본 연구에서는 JDI에서 주로 사용되는 대표적인 6개 요인들을 기준으로 분석하였다. 이를 통해 각 산업 별로 응답자들이 어떠한 직무만족과 관련된 토픽에 대하여 이야기하는 지 알 수 있고 토픽들의 확률을 통하여 직무만족요인의 비율이 어떻게 분포하는지 분석할 수 있을 것으로 기대된다.

많은 연구에서 토픽모델링 수행 후 단어들 간의 구조적 유사성을 탐색적으로 파악하기 위하여 네트워크 분석을 주로 실시했는데, Chae et al.(2015)는 토픽에 출현하는 단어들이 어떠한 관계성을 가지면서 출현하는지 알아보기 위해 네트워크 분석의 일종인 동시출현 단어분석을 실시했다. 이에 본 연구에서는 토픽들을 구성하는 단어들의 관계, 구조를 탐색하기 위하여 동시출현 단어분석을 실시할 것이다. 이를 통해 각 직무만족요인에서 단어들이 어떠한 방식으로 출현하는지 살펴보고 요인들에 대하여 심층적으로 이해할 것이다.

4. 연구 분석 및 결과

4.1 데이터 수집

기업 리뷰는 R을 이용한 웹 크롤링(Web craw-

ling) 기법을 이용하여 잡플래닛에 올라온 산업 군별 110개의 회사에 대한 리뷰를 수집하였다. 리뷰는 잡플래닛 서비스를 실시한 2014년 4월부터 2016년 1월 14일까지 생성된 모든 리뷰 데이터를 수집하였다. 수집된 리뷰 결과는 <Table 1>과 같다. 리뷰는 총 만족도를 포함한 총 6가지 종류에 대한 직무만족 점수와 기업에 대한 총평, 기업의 장점, 단점에 대한 서술로 구성되어 있다. 이 중 본 연구에서는 기업에 대한 총평을 이용하여 직무만족분석을 실시하였다.

<Table 1> Number of Job Reviews by Industries

Industry	Job review
S	4,809
MC	14,190
MP	3,048
DTT	6,180
E	3,713
A	2,754
IT	11,482
MD	4,544
F	5,128
OA	6,181
Total	62,029

*The acronyms used in this table refer to section 2.4.

4.2 데이터 전처리

수집된 리뷰 데이터는 R에서 제공하는 텍스트 전 처리와 관련된 KoNLP, tm, stringr 패키지를 통해 명사, 형용사, 동사로 추출되어 Term-document Matrix로 구축되었다. 이 때 리뷰 작성자의 띄어쓰기 오류로 인해 의존명사, 조사 등이 명사로 인식되어 제거가 되지 않은 채로 Term-document Matrix에 포함되는 경우가 있었다. 예를 들어, ‘통신뿐만 아니라 IT, 미디어 등 회사 내에서 다양한 직무를 수행할 수 있어 ‘T’자형 인재로 회사 내에서 커리어패스를 쌓을 수 있음’과 같은 리뷰를 보면, ‘뿐만’은 의존명사로써 명사에 붙여 쓰

는 것을 원칙으로 하는 국어 문법이 있다. 하지만 위와 같이 띄어쓰기 실수를 할 경우 KoNLP는 패키지가 띄어쓰기를 기반으로 단어를 구분하여 제거하기 때문에 ‘뿐만’이 독립명사로 인식이 되어 제거가 되지 않는 경우가 발생하곤 한다.

이를 해결하기 위해 본 연구에서는 의존명사, 조사, 접속사를 비롯하여 ‘뿐만아니라’와 같은 띄어쓰기가 이루어지지 않아서 완전히 새로운 명사로 인식되는 542개의 단어들을 불용어(Stopwords) 사전으로 구축하였다. 이렇게 구축된 불용어 사전은 의미 없는 단어를 제거하는 텍스트 전 처리에 활용되었다.

4.3 LDA 토픽모델링 분석 결과

앞서 전 처리한 데이터를 R에서 lda 패키지를 이용하여 10개의 산업별로 각각 LDA 토픽모델링 분석을 실시하고 총 30개의 토픽으로 분류했다. 다음의 Table 2는 서비스업의 리뷰 데이터를 이

<Table 2> Result of LDA Topic Modeling(Service)

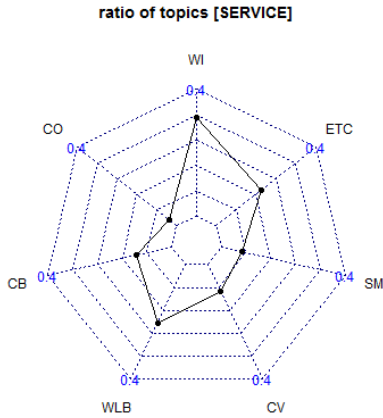
SM (0.032)	WI (0.032)	CB (0.039)	ETC (0.028)	CV (0.028)	WLB (0.041)	...
회사 0.067	서비스 0.059	회사 0.044	회사 0.051	회사 0.061	복지 0.057	...
직원들 0.045	다양한 0.042	체계적 0.036	성장 0.041	기업 0.033	연봉 0.05	...
복지 0.037	사람 0.041	시스템 0.032	기업 0.039	보수적 0.03	비하 0.041	...
성장 0.027	경험 0.039	복지 0.025	사업 0.031	분위기 0.03	급여 0.036	...
분위기 0.021	배우 0.036	동종업계 0.024	안정적 0.029	문화 0.027	대기업 0.032	...
생각 0.020	만나 0.031	대기업 0.023	매출 0.021	기업문화 0.021	회사 0.028	...
기업 0.019	사람들 0.029	체계 0.023	지속적 0.018	조직문화 0.019	괜찮 0.026	...
직원 0.019	추천 0.023	비하 0.021	상황 0.013	오너 0.018	업무강도 0.025	...
경영진 0.017	마인드 0.018	대하 0.02	새롭 0.013	자유로운 0.013	동종업계 0.02	...
마인드 0.016	서비스업 0.014	잡히 0.02	경제 0.012	군대식 0.012	분위기 0.018	...

*The acronyms used in this table refer to section 2.3.

용한 LDA 토픽모델링 분석 결과 중 일부를 보여 주고 있다.

<Table 2>에서 각각의 열은 전체 문서에서 출현하는 토픽을 의미한다. 또한 각각의 열 내부에 있는 단어들은 해당 토픽에서 출현하는 단어들을 나타낸다. 예를 들면 다섯 번째 칼럼은 사내문화라는 직무만족요인으로 재분류된 토픽이고, 각각 ‘회사’, ‘기업’, ‘보수적’, ‘분위기’와 같은 단어들은 사내문화라는 토픽에 출현하는 단어들을 빈도가 높은 순으로 나열한 것이다. 또한 토픽에 할당된 숫자와 단어에 할당된 숫자는 각각 전체 문서에서 해당 토픽이 출현할 확률, 해당 토픽에서 각 단어들이 출현할 확률을 나타낸다. 예를 들면 다섯 번째 토픽의 경우 전체 문서에서 2.8%의 확률로 출현하고 ‘회사’라는 단어는 이 토픽에서 6.1%의 확률로 출현한다는 점을 나타낸다.

위와 같이 분류된 토픽들은 각 토픽에 출현할 확률이 높은 단어들을 기준으로 분류하였다. 예를 들면 다섯 번째 칼럼의 경우 ‘보수적’, ‘분위기’, ‘문화’, ‘기업문화’, ‘조직문화’, ‘자유로운’, ‘군대식’이라는 사내문화와 관련된 단어들의 출현확률이 높게 나타났다. 이를 통해 해당 토픽을 사내문화와 관련된 요인으로 분류하였다. 하지만 몇몇 토픽의 경우 토픽에 출현할 확률이 높은 단어뿐만 아니라 이들 간의 문맥적 의미도 고려하여 추출해야 한다. 예를 들면 여섯 번째 칼럼의 경우 ‘복지’, ‘연봉’, ‘급여’와 같은 급여 및 복지와 관련된 단어들의 출현 확률이 높다. 하지만 세 번째 행을 보면 ‘비하’라는 단어에서 전 처리된 ‘비하’라는 단어의 출현 확률이 높다. 이와 같이 아래의 ‘업무강도’, ‘분위기’와 같은 업무와 삶의 균형에 관한 단어들이 높게 출현한다. 이 경우 텍스트 전 처리 전의 리뷰(raw data)에서 이러한 단어들이 어떻게 구성되어 있는지 파악하여 토픽을 재분류 하였다. 해당 예시에서는 ‘급여에 비하여 업무강도가 괜찮다.’와 같은 업무와 삶의 균형과 관련된 문장들이 주로 나타나 있는 점을 활용하여 급여 및 복지가 아닌 업무와 삶의 균형으로 분류하였다.



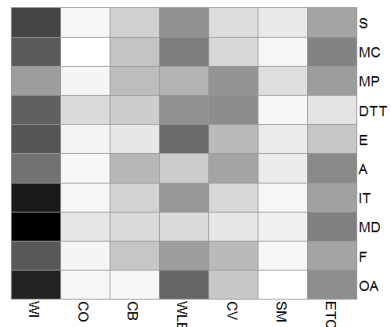
<Figure 3> Ratio of Job Satisfaction Factors(Service)

<Figure 3>을 보면 기타요인을 제외한 직무만족요인 중 직무내용(WI)이 가장 높은 비율을 차지하고 그 다음으로 일과 삶의 조화(WLB), 급여 및 복지(CB), 사내문화(CV)가 중간 비율을 차지함을 볼 수 있다. 또한 승진기회 및 가능성(CO)과 경영진(SM)에 대한 비율은 상당히 낮은 것을 볼 수 있다.

<Table 3>은 위와 같은 방식으로 각각의 산업에 대해 직무만족 요인으로 분류한 결과로서, 각 산업간 직무 만족 요인에 대해 비율을 통해 보여주고 있다. <Figure 4>는 이러한 비율을 히트맵을 통해 시각화하여 보여주고 있다. 히트맵은 색이 진할 수록 더 높은 비율을 표현한다.

<Figure 4>의 히트맵에서 직무만족요인을 기준으로 살펴보면 전체 산업에서 직무내용의 비율이 가장 높음을 알 수 있다. 이를 통해 산업의 종류에 관

계없이 근로자들은 리뷰에서 직업만족요인 중 직무내용 자체에 대해서 가장 많이 말한다는 점을 알 수 있다. 실제 직무만족도 이론에서 직무만족도에 영향을 미치는 가장 큰 요인이 직무내용이라는 선행 연구와 비교할 때 큰 차이가 없는 것을 볼 수 있다. 또한 <Figure 4>에서 직무만족요인 중 급여와 복지, 일과 삶의 조화, 사내문화의 비율이 기관/협회를 제외한 다른 모든 산업에서 비슷한 형태로 나타나는 것을 볼 수 있다. 뿐만 아니라 승진기회 및 가능성과 경영진에 관한 요인은 전체 산업에 걸쳐서 거의 나타나지 않는 점을 볼 수 있다. 이를 통해 산업간 직무만족요인의 비율이 비슷한 분포를 이루는 사실을 알 수 있다. 하지만 위의 결과만으로 산업 간 직무만족요인이 왜 비슷한 패턴이 발생하는지 알기 어렵다. 그러므로 본 연구는 동시출현 단어분석을 통해 이러한 패턴이 발생하는 이유를 살펴볼 것이다.



<Figure 4> Heatmap for Ratio of Job Satisfaction Factors by Industries

<Table 3> Ratio of Job Satisfaction Factors by Industries

	WI	CO	CB	WLB	CV	SM	ETC
S	0.311	0.031	0.114	0.204	0.093	0.067	0.180
MC	0.283	0.000	0.137	0.229	0.099	0.028	0.224
MP	0.187	0.032	0.150	0.160	0.198	0.087	0.187
DTT	0.233	0.074	0.080	0.240	0.174	0.060	0.138
E	0.290	0.038	0.070	0.257	0.151	0.062	0.132
A	0.247	0.024	0.157	0.119	0.178	0.062	0.214
IT	0.357	0.026	0.107	0.197	0.101	0.028	0.182
MD	0.394	0.068	0.097	0.098	0.067	0.050	0.226
F	0.286	0.035	0.132	0.188	0.151	0.030	0.179
OA	0.344	0.030	0.027	0.266	0.127	0.000	0.206

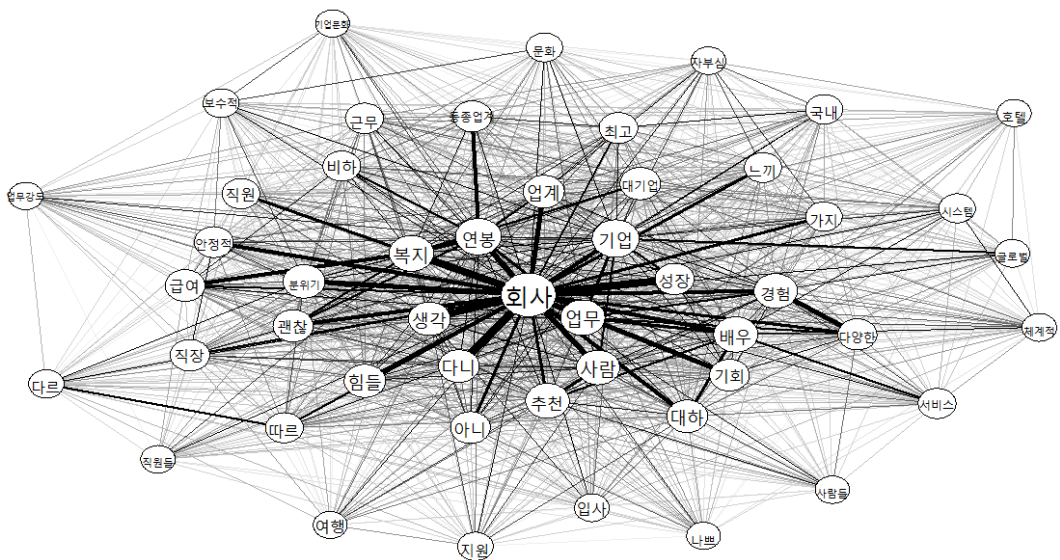
4.4 동시출현 단어분석 결과

본 연구에서는 왜 LDA 토픽모델링 결과 산업 간 직무만족요인의 차이가 존재하지 않았는지 살펴보기 위해 산업별 동시출현 단어분석을 실시했다. 동시출현분석은 각 산업별 출현 빈도가 높은 상위 50개의 단어에 대하여 실시했다. 그 결과 모든 산업에 걸쳐서 직무만족요인 중 직무내용(WI)에 해당하는 단어끼리 동시출현빈도가 높게 나타났고 급여와 복지(CB), 일과 삶의 조화(WLB), 사내문화(CV)에 해당하는 단어들끼리 동시출현빈도가 높게 나타났다. 이러한 현상은 왜 LDA 토픽모델링에서 직무내용이 다른 요인에 비해 독보적으로 높게 나타나고 급여와 복지, 일과 삶의 조화, 사내문화에 관한 요인이 비슷하게 나타났는지 설명해준다. 즉 리뷰 응답자들은 직무만족과 관련하여 주로 직무내용에 대해서 이야기한다. 특히, 직무내용에 관하여 이야기할 때 오직 직무내용과 관련된 것만 이야기한다. 하지만 급여와 복지, 일과 삶의 조화, 사내문화에 대해서 말할 경우 직무내용을 이야기할 때와 달리 위 세 요인들에 관하여 함께 이야기하기 때문에 산업별로 이 요인들의 비

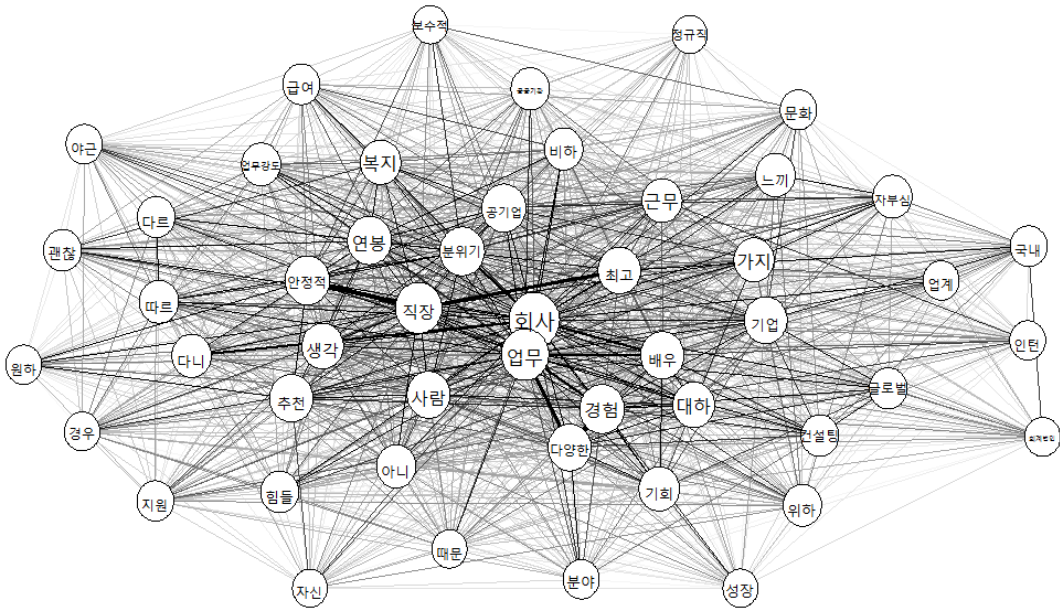
율이 유사하게 나타났던 것이다.

<Figure 5>는 서비스업에 대한 동시출현 단어 분석의 그래프이다. 그래프의 우측 하단을 보면 ‘성장’, ‘경험’, ‘다양한’, ‘업무’, ‘배우’, ‘기회’와 같은 직무내용과 관련된 단어들이 가까이 위치하고 있는 것을 볼 수 있다. 이를 통해 직무내용에 대한 단어는 이 요인에 대한 단어끼리만 동시출현 빈도가 높은 걸 알 수 있다. 반면 좌측 상단을 보면 ‘연봉’, ‘복지’, ‘급여’와 같은 급여와 복지를 나타내는 단어, ‘힘들’, ‘업무강도’, ‘안정적’과 같은 일과 삶의 조화에 대한 단어와 ‘문화’, ‘보수적’, ‘분위기’, ‘기업문화’와 같은 사내문화와 관련된 단어들이 서로 가까이 위치하는 것을 볼 수 있다. 이를 통해 급여와 복지, 일과 삶의 조화, 사내문화에 대해서 말할 경우 주로 같이 말하는 것을 알 수 있다.

<Figure 6>은 기관/협회에 관한 동시출현 단어 분석에 대해 보여주고 있다. LDA 토픽모델링 결과 기관/협회의 직무만족요인의 분포가 약간 다르게 나타났는데, 동시출현 단어분석에서도 이러한 차이가 있는 지를 살펴보았으나 차이가 존재하지 않는 것으로 드러났다. <Figure 6>에서 나타난 결과를 살펴보면 <Figure 5>의 서비스업에서 본 것



<Figure 5> Co-occurrence Network for Service



<Figure 6> Co-occurrence Network for Organization/Association

과 비슷하게 우측하단에 직무내용과 관련된 단어들이, 좌측 상단에는 급여와 복지, 일과 삶의 조화, 사내문화에 관련된 단어들이 주로 분포하는 것을 볼 수 있다. 다만 좌측 상단 부분에 ‘안정적’, ‘직장’, ‘최고’라는 일과 삶의 조화와 관련된 단어들의 동시출현빈도가 상당히 높은 것을 볼 수 있다. 이 때문에 <Figure 4>에서 기관/협회의 일과 삶의 조화와 관련된 요인이 다른 산업에 비해 유달리 높게 나타난 것을 알 수 있다. 이처럼 기관/협회의 경우 약간의 특이성이 관찰되지만 전체적인 직무만족요인과 관련된 단어들이 분포하는 양상은 다른 산업과 유사하다고 판단할 수 있다.

5. 결 론

본 연구는 산업별 직무만족요인의 차이가 있는 지 살펴보기 위하여 기업 리뷰 데이터를 이용한 LDA 토픽모델링과 동시출현 단어분석을 통해 비교분석을 실시하였다. 우선 LDA 토픽모델링을 통해 살펴본 결과 기업 리뷰에서 6개의 직무만족요인 중 직무내용, 급여 및 복지, 업무와 삶의 균형,

사내문화와 관련된 요인이 주로 나타났다. 이 중 직무내용과 관련된 요인의 비율은 전체 산업에서 공통적으로 가장 높게 나타났고 급여 및 복지, 업무와 삶의 균형, 사내문화의 비율은 직무내용에 비해 상대적으로 낮게 나타났지만 산업간에 유사한 분포가 나타났다. 그리고, 동시출현 단어분석을 이용하여 기업 리뷰 데이터를 심층적으로 분석한 결과 급여 및 복지, 업무와 삶의 균형, 사내문화와 관련된 단어들이 동시출현빈도가 높게 나타났다. 이를 통해 산업 간의 직무만족요인의 분포가 유사하게 나타남을 확인했다.

이러한 연구 결과는 사후 산업간 직무만족요인에 관한 후속 연구를 수행하는 데 도움이 될 것으로 기대된다. 더 나아가 실무에서도 본 연구의 방법론을 활용한다면 각 기업들은 리뷰 데이터를 통해 직원들의 직무에 대한 만족도를 좀 더 체계적으로 분석할 수 있을 것으로 판단된다. 또한 본 연구는 기존의 조직행동연구에서 활용되지 않았던 기업 리뷰 데이터의 활용 가능성을 제시하고 있다.

특히, 최근 SNS가 기하급수적으로 발달하고 스마트폰이 보편화됨에 따라서 기업 리뷰 데이터와

같은 UGC가 급속도로 증가할 추세를 보이기 때문에 본 연구 결과는 이를 활용한 연구의 하나의 사례 역할을 할 수 있을 것으로 기대된다.

하지만 본 연구는 다음과 같은 한계점이 있다. 우선 산업별 직무만족요인에 대한 비교 분석을 실시했지만 분석 결과가 실제 리뷰 데이터를 통해 살펴본 직무만족도에 유의미한 영향을 미치는가에 대한 검증은 이루어지지 않았다. 향후 연구에서는 직무만족요인의 비율 차이에 따라 실제 직무만족도에 대한 영향력이 차이가 발생하는지 분석을 수행해야 될 것이다. 또한 본 연구에서는 텍스트 전처리 과정에서 명사, 형용사와 동사만 추출하여 문법적인 오류를 최소화하고 기업과 회사와 같은 의미적으로 같은 단어를 통일하려 했으나 기술적 한계로 인해 이러한 전처리 과정이 완벽하게 이루어지지 않았다. 그러므로 연구 결과의 정확성을 높이기 위해 기본형 판별을 통해 개선해야 될 것으로 보인다. 이러한 한계점에도 불구하고 본 연구는 조직행동연구에서 리뷰 데이터의 활용 가능성을 보여주었다는 점에서 시사점을 가질 것으로 기대된다.

References

- Blei, D.M., "Probabilistic Topic Models", *Communications of the ACM*, Vol.55, No.4, 2012, 77-84.
- Blei, D.M., A.Y. Ng, and M.I. Jordan, "Latent Dirichlet Allocation", *the Journal of machine Learning research*, Vol.3, 2003, 993-1022.
- Bussing, A., T. Bissels, V. Fuchs, and K.M. Perrari, "A Dynamic Model of Work Satisfaction : Qualitative Approaches", *Human Relations*, Vol.52, No.8, 1999, 999-1028.
- Callon, M., J.P. Courtial, W.A. Turner, and S. Bauin, "From Translations to Problematic Networks : An Introduction to Co-Word Analysis", *Social science information*, Vol. 22, No.2, 1983, 191-235.
- Cha, Y.J., J.H. Lee, J.E. Choi, and H.W. Kim, "A Topic Modeling Approach to Marketing Strategies for Smartphone Companies", *Knowledge Management Research*, Vol.16, No.4, 2015, 69-87.
- (차윤정, 이지혜, 최지은, 김희웅, "연구논문 : 소셜 미디어 토픽모델링을 통한 스마트폰 마케팅 전략 수립 지원", *지식경영연구*, 제16권, 제4호, 2015, 69-87.)
- Chae, S.H., J.I. Lim, and J.Y. Kang, "A Comparative Analysis of Social Commerce and Open Market Using User Reviews in Korean Mobile Commerce", *Journal of Intelligence and Information Systems*, Vol.21, No.4, 2015, 53-77.
- (채승훈, 임재익, 강주영, "사용자 리뷰를 통한 소셜커머스와 오픈마켓의 이용경험 비교분석", *지능정보연구*, 제21권, 제4호, 2015, 53-77.)
- Cho, K.H. and J.B. Ju, "Comparing Work Satisfaction among Different Types of Organizations : An Application of Publicness Model", *Korean Public Administration Review*, Vol.35, No.3, 2001, 163-182.
- (조경호, 주재복, "조직유형별 직업만족도 비교연구", *한국행정학보*, 제35권, 제3호, 2001, 163-180.)
- Clark, A.E., "Job Satisfaction in Britain", *British journal of industrial relations*, Vol.34, No.2, 1996, 189-217.
- Coughlan, M., P. Cronin, and F. Ryan, "Survey Research : Process and Limitations", *International Journal of Therapy and Rehabilitation*, Vol.16, No.1, 2009.
- Cranny, C., P.C. Smith, and E. Stone, "Job Satisfaction : How People Feel About Their Jobs", 1992.
- Decay, S.H., "Peering through Glassdoor, Com What Social Media Can Tell Us About Em-

- ployee Satisfaction And”, *Proceedings of Conference on Corporate Communication 2013*, 2013.
- Demaio, T.J., “Social Desirability and Survey”, *Surveying Subjective Phenomena*, Vol.2, 1984, 257.
- Doopedia, “Ucc(User Created Content)”, available at <http://terms.naver.com/entry.nhn?docId=1255390&cid=40942&categoryId=32854>, 2016.
- Grimmer, J., “A Bayesian Hierarchical Topic Model for Political Texts : Measuring Expressed Agendas in Senate Press Releases”, *Political Analysis*, Vol.18, No.1, 2010, 1-35.
- Herzberg, F., “The Motivation-Hygiene Concept and Problems of Manpower”, *Behavioral concepts in management*, 1968, 35.
- Jin, S.A., C.E. Heo, Y.K. Jeong, and M. Song, “Topic-Network Based Topic Shift Detection on Twitter”, *Journal of the Korean Society for Information Management*, Vol. 30, No.1, 2013, 285-302.
- (진설아, 허고은, 정유경, 송 민, “트위터 데이터를 이용한 네트워크 기반 토픽 변화 추적 연구”, *정보관리학회지*, 제30권, 제1호, 2013, 285-302.)
- Jung, J.M. and Y.J. Kim, “Effects of Person-Organization Fit and Job Satisfaction on Organizational Commitment and Turnover Intention—a Case of Media Firms’ Employee”, *Korean Journal of Broadcasting and Telecommunication Studies*, Vol.22, No.3, 2008, 290-331.
- (정재민, 김영주, “미디어 기업 종사자의 조직적합도와 직업만족도가 조직성과에 미치는 영향”, *한국방송학보*, 제22권, 제3호, 2008, 290-331.)
- Kim, D.S., W. William Xiu Shun, M.S. Lim, C. Liu, N.G. Kim, J.H. Park, W.Y. Kil, and H.S. Yoon, “A Methodology for Analyzing Public Opinion About Science and Technology Issues Using Text Analysis”, *Journal of Information Technology Services*, Vol.14, No.3, 2015a, 33-48.
- (김다솜, W. William Xiu Shun, 임명수, 류 신, 김남규, 박준형, 길우영, 윤한술, “텍스트 분석을 활용한 과학기술이슈 여론 분석 방법론”, *한국IT서비스학회지*, 제14권, 제3호, 2015a, 33-48.)
- Kim, H.J. and M. Song, “A Study on the Research Trends in Domestic/International Information Science Articles by Co-Word Analysis”, *Journal of the Korean Society for Information Management*, Vol.31, No.1, 2014, 99-118.
- (김하진, 송 민, “동시출현단어 분석을 통한 국내외 정보학 학회지 연구동향 파악”, *정보관리학회지*, 제31권, 제1호, 2014, 99-118.)
- Kim, J.H., Y.E. Song, Y.S. Jin, and O.B. Kwon, “Applying Text Mining to Identify Factors Which Affect Likes and Dislikes of Online News Comments”, *Journal of Information Technology Services*, Vol.14, No.2, 2015b, 159-176.
- (김정훈, 송영은, 진윤선, 권오병, “텍스트마이닝을 통한 댓글의 공감도 및 비공감도에 영향을 미치는 댓글의 특성 연구”, *한국IT서비스학회지*, 제14권, 제2호, 2015, 159-176.)
- Kinicki, A.J., F.M. Mckee-Ryan, C.A. Schriesheim, and K.P. Carson, “Assessing the Construct Validity of the Job Descriptive Index : A Review and Meta-Analysis”, *Journal of Applied Psychology*, Vol.87, No.1, 2002, 14.
- Lakin, T., “Why You Shouldn’t Ignore Glassdoor”, HR, 2015.
- Lee, H.J., “A Study on Classifications of Useful Customer Reviews by Applying Text Mining Approach”, *Journal of Information Tech-*

- nology Services*, Vol.14, No.4, 2015, 159-169.
 (이홍주, “텍스트 마이닝을 활용한 고객 리뷰의 유용성 지수 개선에 관한 연구”, *한국IT서비스학회지*, 제14권, 제4호, 2015, 159-169.)
- Locke, E.A., “The Nature and Causes of Job Satisfaction”, *Handbook of industrial and organizational psychology*, Vol.1, 1976, 1297-1343.
- Park, J.G., Y.J. Kim, and J.M. Jung, “Impacts of Creative Organizational Culture on Job Satisfaction of Media Firms’ Employees-Perception on the Creative Work Environment”, *Korean Journal of Broadcasting and Telecommunication Studies*, Vol.23, No.1, 2009, 169-207.
 (박종구, 김영주, 정재민, “미디어기업의 창의적 조직문화가 직무만족도에 미치는 영향”, *한국방송학보*, 제23권, 제1호, 2009, 169-207.)
- Park, J.H. and M. Song, “A Study on the Research Trends in Library and Information Science in Korea Using Topic Modeling”, *Journal of the Korean Society for Information Management*, Vol.30, No.1, 2013, 7-32.
 (박자현, 송민, “토픽모델링을 활용한 국내 문헌정보학 연구동향 분석”, *정보관리학회지*, 제30권, 제1호, 2013, 7-32.)
- Ricketta, M., “The Causal Relation between Job Attitudes and Performance : A Meta-Analysis of Panel Studies”, *Journal of Applied Psychology*, Vol.93, No.2, 2008, 472.
- Scarpello, V. and J.P. Campbell, “Job Satisfaction : Are All the Parts There?”, *Personnel psychology*, Vol.36, No.3, 1983, 577-600.
- Zhao, W.X., J. Jiang, J. Weng, J. He, E.P. Lim, H. Yan, and X. Li, “Comparing Twitter and Traditional Media Using Topic Models”, 2011, In *Advances in Information Retrieval*, Springer, ISBN 3642201601.

◆ About the Authors ◆



Dongwook Kim (kdw4521@gmail.com)

Dongwook Kim is a undergraduate student in e-Business at Ajou University. His current research interests include Big data, CRM and Text mining.



Juyoung Kang (jykang@ajou.ac.kr)

Professor Juyoung Kang is currently a Full Professor of e-Business at School of Business, Ajou University. She received her Ph.D. in Management Engineering from Korea Advanced Institute of Science and Technology (KAIST) in 2005. She has more than 50 refereed publications in academic journals and conference proceedings and has developed Intelligent Systems and E-Commerce applications with various industrial partners. Her current research interests are in the fields of text mining, cloud computing, big data, and intelligent systems. and etc.



Jay Ick Lim (limjay@ajou.ac.kr)

Professor Jay Ick Lim is Associate Professor in the department of e-Business at School of Business, Ajou University. He got his B.A. degree in economics from Seoul National University, and received his Ph.D. in Management Information Systems from the University of Iowa in 1990. He taught at Cleveland State University and Rutgers University, USA. His current research interests include IT service, innovative business models, e-business, intelligent information & knowledge management, and etc.