

비정형 데이터분석과 시각화 “한국심리학회지: 산업 및 조직(2010년~2017년)”을 중심으로

이 인 선[†]

영남대학교

나 은 영[‡]

안동대학교

빅데이터가 보급되면서 비정형자료 분석기술에 관한 관심이 증가하고 있으며, 방대한 양의 학술 데이터에서 가치 있는 정보를 도출해 내는 연구들이 수행되고 있다. 대부분 대용량 텍스트를 인간이 수동적으로 분석하는 것은 시간과 노력이 많이 들고 어려우므로 이를 보완하기 위해 자동으로 분류해줄 수 있는 기술이 필요하다. 이에 본 연구는 2010년부터 2017년까지 발간한 ‘한국심리학회지: 산업 및 조직’ 논문의 주요어, 논문제목, 초록과 같은 비정형자료를 이용하여 분석하고 시각화하였다. 연구결과는 다음과 같다. 첫째, 주요어를 워드클라우드로 시각화한 결과 ‘조직몰입’이 빈번히 사용되었으며, 직무만족, 직무열의, 이직의도, 정서노동 순으로 나타났다. 둘째, 논문제목과 초록을 LDA 확률에 근거하여 자동으로 10개의 주제로 분류하였으며, 주제 8(조직/몰입/지각)이 가장 높았고, 주제 5(행동/관리/상사)는 가장 낮았다. 셋째, 주저자와 교신저자와의 관계는 6개의 큰 집단과 다수의 작은 집단으로 나타났으며, 집단 내의 유력자를 알 수 있었다. 본 연구에서는 방대한 양의 학술 데이터에서 보다 빠르고 쉽게 정보를 도출하여 제공함으로써 관련 연구자들이 또 다른 연구에 접근할 수 있도록 도움을 줄 수 있다는데 의의를 둔다.

주요어 : 비정형자료, LDA, 빅데이터, I-O 심리학

[†] 주저자 : 이인선, 영남대학교 심리학과, tksaor21@naver.com

[‡] 교신저자 : 나은영, 안동대학교 기초교육원, ney111@naver.com

4차 산업혁명을 이끄는 신기술의 위협으로 직업의 판도가 바뀔 것이라는 전망은 취업자나 미취업자에게 희망 또는 불안으로 다가오고 있다. 세계 여러 국가들과 기업들은 경쟁력 확보라는 명분으로 4차 산업 관련 기술 개발에 적극적으로 나섰다. 기술의 미래 모습을 양적, 질적, 내용적으로 예측하는 기술 예측을 위해서는 사회적 요소를 고려해야하지만 현재는 기술에 의해 사회가 변화한다는 관점에서 대부분 기술을 중심으로 개발하고 있다. 사회적 요소를 간과한 기술 예측은 실효성이 낮을 수 있다. 특히 우리는 4차 산업혁명의 시작점에 있기 때문에 모든 산업과 기업의 구성원은 개인과 조직에 어떤 영향을 미칠 것인가를 생각해야만 한다. 모든 사람에게 기회를 주는 변화가 되도록 미래를 함께 만들어나가야 하는 것이 우리의 과업이다. 작업장에서 일을 할 때 나타나는 인간의 행동과 정서에 관심을 가지며, 이에 대한 연구를 바탕으로 조직과 구성원의 복지를 향상시키고자 하는 분야인 산업 및 조직(I-O, Industrial-organizational) 심리학은 이러한 변화를 민감하게 받아들여야 한다. 미국 노동부 통계에 따르면 2012년부터 2022년까지 가장 빠르게 성장하는 20개 분야 중 I-O 심리학이 53.4%의 성장률로 1위를 차지했다(Bureau of Labor Statistics, 2013). 이것은 I-O 심리학에 대한 관심이 높아질 것이라는 것을 보여준다. I-O 심리학은 조직이 인적자원들을 효율적으로 관리하고 개발하여 조직의 효과성을 높이는 데 도움을 제공하고, 일을 하는 사람들 각자가 맡은 일에 만족하고, 일의 의미를 찾을 수 있도록 도움으로써 자아실현을 이루고 삶의 보람을 가질 수 있도록 하는 분야이다(Muchinsky, 2012). 그러므로 우리나라의 I-O 심리학의 연구들을 분석하여 활용할

수 있도록 기초자료를 제공하는 것은 의미 있는 작업일 것이다.

I-O 심리학 연구자들이 변화를 인식하고 학술지 논문과 같은 인적자원의 보고에서 자신의 의견을 피력하거나 가치를 창출하는 것은 중요한 역할을 한다고 할 수 있다. 따라서 학회지 게재 논문의 실태나 현상을 파악하는 동향분석이 필요하다. 유태용과 탁진국(1998)은 '한국심리학회지: 산업 및 조직'에서 학회지를 발간부터 시작하여 이후 10년 동안 게재된 논문을 저자특성, 연구주제, 연구장면, 연구방법, 연구 참가자 유형, 통계분석 종류 및 사용빈도 등을 분석하였고, 연구자들이 직접 평정 절차를 통해 일치되는 정보를 추출하였다. 정재삼, 강민진, 이혜진 및 조지현(2014)은 '기업교육연구', '인적자원관리연구', '한국심리학회지: 산업 및 조직'의 3개 학회지를 분석하였으나 HRD이슈와 동향만을 탐색하였다. 대부분 연구동향 분석논문에서 연구자가 각자 인지적 능력을 사용하여 논문을 연구의 목적에 맞추어서 게재 논문을 분석하였고, 일치하지 않는 부분에 대해서는 연구자들이 합의하는 방식으로 분류하도록 하고 있다.

우리나라는 세계 최고 수준의 초고속 통신망을 구축하고 있는 IT국가이며 이는 학문분야에서도 정보화된 논문들이 학술 커뮤니티를 통해 학자들에게 시시각각 제공되고 있는 상황으로 기술변화에 민첩한 대응이 필요하다. 그러나 자료가 축적될수록 인간의 인지적 능력으로 분석하는 것은 한계가 있다. 연구의 동향을 분석하는 데 소요하는 시간과 노력을 줄여 효과적으로 수행할 수 있는 기술의 활용이 필요하다. 최근 일부 분야에서는 방대한 양의 학술 데이터에서 의미 있는 정보와 지식을 도출해 내는 분석 연구들이 수행되고 있다.

산업 및 조직심리학 분야

I-O 심리학은 ‘선발과 배치’, ‘교육과 개발’, ‘직무수행평가’, ‘조직개발’, ‘작업생활의 질’, ‘인간공학’의 여섯 가지 분야로 나누었다 (Muchinsky, 2012). 선발과 배치는 직책에 따른 업무의 특징과 필요역량은 무엇인가, 직원의 선발, 배치에 필요한 역량 측정은 어떻게 할 것인가, 각자의 적성과 흥미에 적합한 직무는 무엇인가 등에 관심을 갖는 것이다. 교육과 개발은 조직원들의 역량을 어떻게 향상시킬 수 있는가, 조직에 필요한 교육 및 개발 프로그램은 무엇인가, 그런 프로그램의 구성과 평가는 어떻게 할 것인가 등에 관심을 갖는 것이다. 직무수행평가는 조직원들이 업무를 얼마나 잘 수행하고 있는가를 평가하기 위한 기준과 원칙은 어떻게 설정할 것인가, 직무수행이 조직에 기여하는 효용성이나 가치는 무엇인가, 작업팀, 조직 내 부서, 조직 자체의 수행이 어느 정도인가 등에 관심을 갖는 것이다. 조직개발은 조직의 구조가 어떻게 구성되어있는가, 조직의 성장을 어떻게 촉진시킬 것인가 조직 내외의 어떤 요소들이 개인의 행동에 영향을 미치는가 등에 관심을 갖는 것이다. 작업생활의 질은 직원들이 일을 함에 있어 큰 의미와 만족감을 느낄 수 있도록 어떻게 해야 하는가, 건강하고 생산적인 일터를 만들기 위해선 어떻게 해야 하는가 등에 관심을 갖는 것이다. 인간공학은 다른 학문들과 관련된 학제 간으로 연구해야 할 분야로, 작업현장의 장비와 기계들을 어떻게 설계해야 사용하기 편리한가, 이를 통해 어떻게 작업환경을 구성해야 직원들이 편리하고 효율적으로 일을 할 수 있는가 등에 관심을 갖는 것이다.

미국 산업 및 조직심리학 10대 트렌드

미국 산업 및 조직심리학회는 2014년부터 매년 작업장에서 해결해야 할 것으로 예측하는 사항에 대한 트렌드 10개를 발표함으로써 산업 조직 심리학자가 인재 관리, 코칭, 평가, 선택, 훈련, 조직 개발, 성과 및 일과 삶의 균형을 포함하여 비즈니스와 관련성이 매우 중요한 작업장 문제를 연구하도록 하고 있다. 미국 산업 및 조직심리학회에서 조사한 트렌드 목록은 산업 및 조직심리학에 관심을 갖고 있으며 연구할 수 있는 역량을 갖춘 연구자들을 대상으로 의견을 수렴하였기 때문에 이 분야의 연구자들에게 연구할 수 있는 주제를 제시하는 역할을 한다고 할 수 있다.

2014년에 선정된 10개의 트렌드는 다음과 같다(Society for Industrial and Organizational Psychology, 2014). 1. 빅 데이터(Big data), 2. 효율성 증대(Increasing efficiency), 3. 원하는 재능을 가진 인재 채용 방법(The talent question), 4. 새로운 테스트 방법(New ways to test), 5. 게이미피케이션(Gamification) 6. 직장에서의 기술 통합(Integration of technology in the workplace), 7. 일과 삶의 균형(Work-life balance), 8. 소셜 미디어(Social media), 9. 원격 근무(Tele work), 10. 정규직에 대한 대안(Alternatives to full-time).

2015년 미국과 캐나다 I-O 심리학자들이 선정한 10개의 트렌드는 다음과 같다(Canadian Psychological Association, 2015). 1. 모바일 평가(Mobile Assessments 2. HR 분석 및 빅 데이터의 지속적인 사용(Continued Use of HR Analytics and Big Data) 3. 일과 삶의 통합(Integration of Work and Nonwork Life) 4. 업무 수행 방식에 대한 기술의 시사점 증가(Increasing Implications of Technology for How Work is Performed) 5. 조

직은 “적은 비용으로 더 많은 것을 수행”을 계속할 것입니다(Organizations Will Continue to “Do More with Less”) 6. 다세대 노동력 관리의 필요성 증가(Increased Need to Manage a Multi-Generational Workforce) 7. 채용, 선택 및 잠재력 유지에 관한 강조(Emphasis on Recruiting, Selecting for, and Retaining Potential) 8. 다양한 계획의 변화 모습(Changing Face of Diversity Initiatives) 9. 기업의 사회적 책임 프로그램의 성장(Growth of Corporate Social Responsibility Programs) 10. 법 변경은 고용 관련 결정에 영향을 미칠 수 있습니다(Changes in Laws May Affect Employment-Related Decisions).

2016년에 선정된 10개의 트렌드는 다음과 같다(Society for Industrial and Organizational Psychology, 2016). 1. 빅 데이터 활용 및 극대화(Leveraging and maximizing big data), 2. 기술 동향(Trends in technology), 3. 가상 팀 관리(Managing virtual teams), 4. 성과 관리 및 개발의 특성 변화(Changing nature of performance management and development), 5. 직원 참여(Employee engagement) 6. 직장에서의 건강과 복지에 대한 관심 확대(Increasing focus on health and wellness in the workplace), 7. 비즈니스 민첩성 및 유연성에 대한 관심 증대(Increased focus on business agility and flexibility), 8. 여러 세대의 일과 삶의 균형(Work-life balance across generations), 9. 건강하고 다양한 인력 구축(Building healthy, diverse workforces), 10. 고용 관련 결정시 소셜 미디어 사용(Using social media to make employment-related decisions).

2017년에 선정된 10개의 트렌드는 다음과 같다(Society for Industrial and Organizational Psychology, 2017). 1. 성과 관리의 특성 변화(The Changing Nature of Performance

Management), 2. 효과적으로 변화에 적응하기(Adapting to Change Effectively), 3. 빅 데이터를 활용하여 데이터 중심 의사 결정(Leveraging Big Data to Make Data-Driven Decisions), 4. 사용자 분석(People Analytics), 5. 작업 방식의 유연성과 그 영향(Flexibility and Its Effect on the Way Work Is Done), 6. 노동력의 특성 변화(The Changing Nature of the Workforce), 7. 직원의 의견 파악(Capturing the Voice of the Employee), 8. 다양성과 포용성의 중요성 증대(Growing Importance of Diversity and Inclusion), 9. 소스, 시스템 및 프로세스 간 데이터 통합(Data Integration Across Sources, Systems, and Processes), 10. 직원 건강 및 복지 향상(Increased Focus on Employee Health and Wellness).

4차 산업혁명 시대는 연구자가 직접 평정하여 분류·분석하는 기존연구 방법도 중요하지만, 이보다 시간과 노력을 감소시켜 빠르게 발전하는 기술에 대응하는 텍스트 마이닝 관련 기술이 필요하다고 할 수 있다.

텍스트 마이닝(Text Mining)

텍스트 형식의 정보를 분석하는 데는 텍스트 마이닝이 필요하다. Kao와 Poteet(2007)는 텍스트 마이닝을 비정형화된 텍스트로부터 흥미롭고 일상적이지 않은 지식을 발견, 추출하는 것이라고 정의하였다. 한국정보화진흥원(2012)은 텍스트 마이닝이 자연어로 구성된 비정형 텍스트 데이터에서 내재된 정보나 패턴 또는 관계를 추출하여 가치와 의미있는 정보를 찾아내는 빅데이터 분석을 위한 마이닝 기법의 하나로 기술하였다. 텍스트 마이닝 기법을 활용한다면, 객관적인 분석이 가능해지고 연구의 주제와 동향을 다각도로 조명해 볼 수

있다. 최근 빅데이터가 보급되면서 대용량 텍스트 분석기술에 관한 관심이 증가하고 있으며 텍스트 마이닝 기술의 중요성이 강조되고 있다. 텍스트 마이닝 기법의 하나인 워드클라우드 등의 기술을 적용하여 산출한 결과물을 시각화할 수 있다. 이는 하나의 테스트에 출현하는 단어의 빈도수에 비례하여 크기로 표출하는 그래프를 표현함으로써 경제적이고 효과적인 분석결과를 제공하여 관심 연구의 관점을 시각적으로 보다 쉽게 전달할 수 있는 기술이다.

Blei, Ng & Jordan(2003)에 의해 제안된 LDA(Latent Dirichlet allocation)는 계층적 베이저안 모델로서 문서 내의 숨겨져 있는 주제들을 찾아내는 알고리즘이다. 여기서 LDA에서 문서는 특정 확률에 의해 선택된 단어들로 구성된 주제들의 혼합 집합으로 표현된다. 이를 분석함으로써 해당 자료가 함께 다루고 있는 주제가 무엇인지를 예측할 수 있으며, 각 주제에 포함될 확률들을 알아낼 수 있다. Blei(2012)는 Science저널과 Yale Law저널의 논문을 대상으로 LDA모델을 활용하여 주제 모델을 도출한 사례를 소개하였으며, 심리학에 활용될 수 있음을 언급하였다. 워드클라우드를 텍스트를 중요도에 따라 시각적으로 나타내고, LDA는 숨어있는 주제를 찾아내고 그 주제에 따라 논문을 확률에 의해 자동으로 분류하는 기술이다.

지금까지 텍스트마이닝의 워드클라우드와 LDA 확률 기법을 활용하여 학술지 연구 분석을 수행한 논문은 찾아보기가 어려웠다. 이에 본 연구에서는 “한국심리학회지: 산업 및 조직” 논문들에 대한 비정형자료 분석을 통해 연구자들에게 새로운 접근에 대한 자료를 제공하고자 한다.

방 법

연구대상 및 자료수집

본 논문은 2010년~2017년 “한국심리학회지: 산업 및 조직”의 학술자료의 주요어, 논문 제목, 초록, 저자와 같은 비정형자료를 이용하여 244편의 논문을 분석하였다.

분석기법

분석방법은 텍스트마이닝 기법의 하나인 워드클라우드를 활용하여 누구나 직관적으로 파악할 수 있도록 핵심단어를 시각화하였으며, 통계적인 추론 기반의 잠재 디리클레 할당(LDA)을 사용하여 연구의 주제를 자동으로 분류하여 I-O 심리학 분야의 주제로 설명하였으며, 주저자와 교신저자의 관계를 시각화하여 몇 개의 집단으로 구성되며 해당 집단의 유력자는 누구인지를 통계패키지 R로 분석하였다.

결 과

자료 분석 결과는 다음의 세 가지 방식으로 나타내었다. 첫째, 비정형자료(주요어)를 이용한 동향 시각화, 둘째, 비정형자료(논문제목과 초록)를 이용한 주제 분류와 시각화, 셋째, 비정형자료(주저자와 교신저자)를 이용한 관계 시각화이다.

비정형자료(주요어)를 이용한 동향 시각화

분석대상이 되는 논문은 2010년~2017년까지의 “한국심리학회지: 산업 및 조직”의 학술

지논문 244편으로 한정하였다. 저자의 의도를 분석하기 위해 저자가 무엇을 주요어로 사용하였는지 워드클라우드를 이용하여 핵심단어를 시각화하였다. 여기서 많이 언급된 논문의 주요어 일수록 크게 표현되었다. 텍스트마이닝 기법은 수치 값의 크기를 비교하는데 의미를 두기보다는 어떤 부분에 더 많은 관심이 있었는지를 중심으로 보아야 한다.

논문의 2010년~2017년까지의 주요어 분석 결과 그림 1과 같이 '조직몰입'이 가장 빈번히 사용되었으며, '직무만족', '직무열의', '이직의도', '정서노동', '직무탈진' 순으로 나타났다. 그 다음으로 '조직시민행동', '변혁적 리더십', '자기효능감', '학습목표지향성' 등에 관심이 있는 것으로 나타났다.

그림 2~그림 9는 동향을 알아보기 위해 연도별로 주요어를 워드클라우드로 시각화하여 저자들이 어떤 부분에 관심이 많았는지를 알아보았다.

2010년에는 '직무열의', '정서노동', 2011년에는 '조직몰입', '직무만족', 2012년에는 '조직시민행동', '매개효과'의 주요어가 빈번히 사용된 것으로 나타났다. 2013년에는 '조직몰입', 2014년에는 '조직몰입', '직무만족', '메타분석', '정서노동', '직무탈진', 2015년에는 '조직몰입', 2016년에는 '직무만족', 2017년에는 '이직의도', '관계갈등', '직무만족'의 단어가 주요어로 빈번히 사용된 것으로 나타났다.

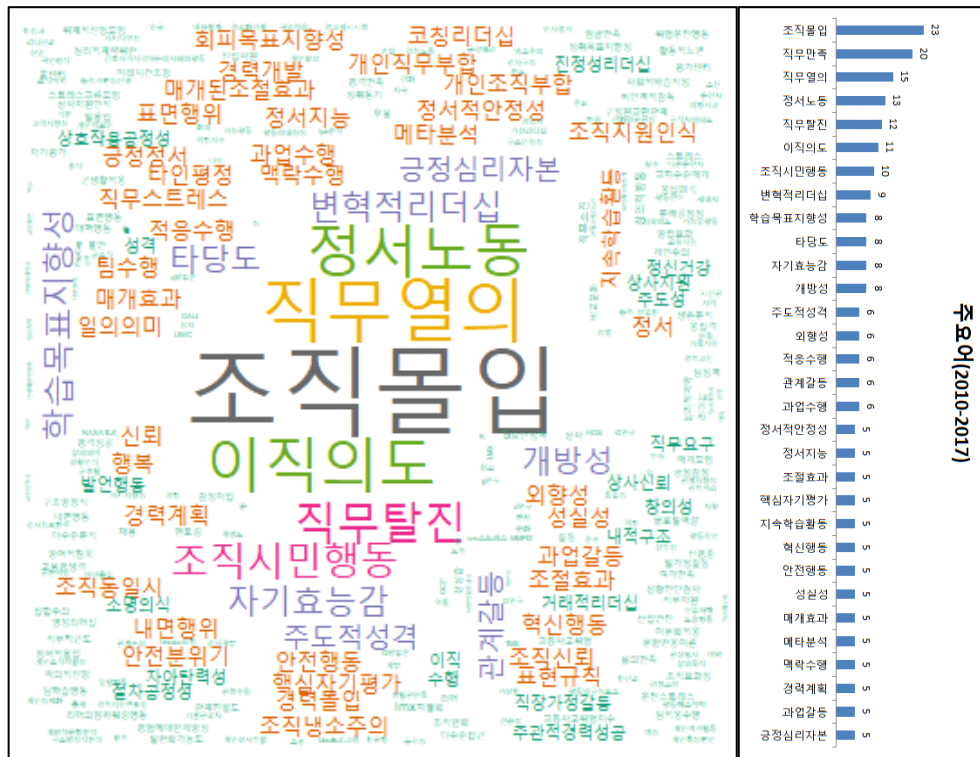


그림 1. 2010-2017년 주요어

2010-2017산업및조직[주제7]

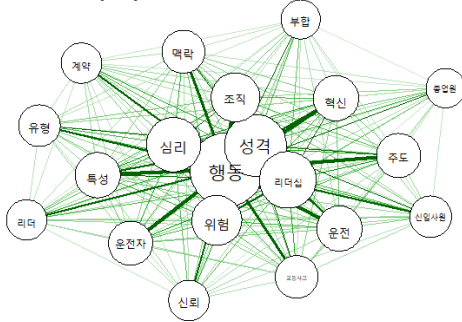


그림 16. LDA 기반 주제7(행동/성격/리더십)

2010-2017산업및조직[주제8]

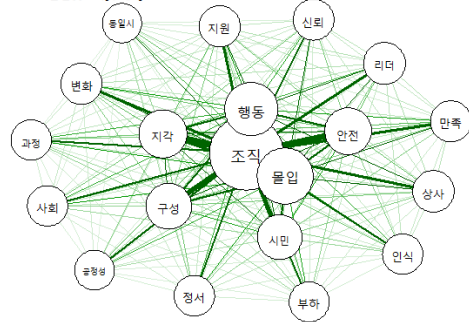


그림 17. LDA 기반 주제8(조직/몰입/지각)

2010-2017산업및조직[주제9]

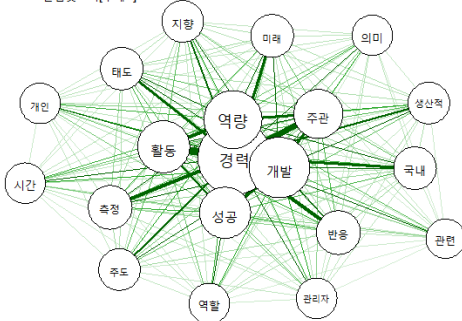


그림 18. LDA 기반 주제9(경력/개발/역량)

2010-2017산업및조직[주제10]

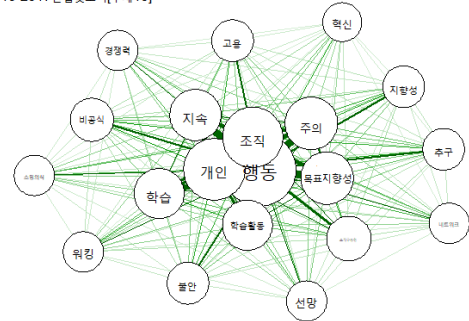


그림 19. LDA 기반 주제10(행동/개인/조직)

주제로 ‘행동, 개인, 조직’을 중심으로 ‘목표지향성, 주의, 지속, 학습, 학습활동’에 관한 단어로 분류되었다.

전체논문을 10개의 주제로 분류하였다. 각 논문은 10개 주제들에 따른 확률의 집합으로 나타났으며, 가장 높은 확률을 기준으로 전체 244편의 논문을 분류한 결과 93.9%를 하나의 주제로 분류하였다. 나머지 4.5%는 두 개의 주제로 묶였으며, 1.6%는 세 개 이상의 주제에 비슷한 확률로 분류되어 결과 분석 시 제외하고 나타내었다. 큰 차이는 없으나, 그림 20과 같이 주제8(조직/몰입/지각)이 14.0%로 가장 많이 분류되었으며, 주제1(직무/만족/심리)

과 주제6(이직/직무)은 각각 12.2%, 주제7(행동/성격/리더십)은 11.8%, 주제2(직무/정서/노동), 주제4(갈등/직장/가정)는 각각 9.6%, 주제9(경력

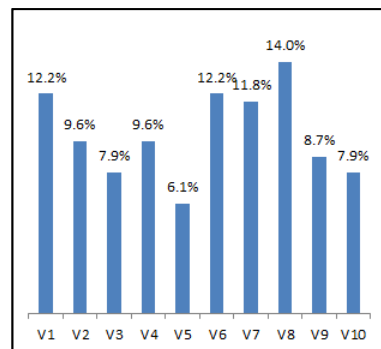


그림 20. 주제별 비율

/개발/역량)는 8.7%, 주제3(자기/적응/탐색)과 주제10(행동/개인/조직)은 각각 7.9%순으로 나타났으며, 주제5(행동/관리/상사)는 6.1%로 다소 낮은 비율로 분류되었다.

비정형자료(주저자와 교신저자)를 이용한 관계 시각화

표 1과 같이 논문의 저자 수가 2인인 경우가 51.68%로 가장 많았으며, 3인(23.8%), 1인(10.7%), 4인(9.8%), 5인(3.7%), 8인(0.4%) 순으로 나타났다. 그림 21과 같이 연도별 1인, 2인, 3

인에 대한 동향을 꺾은선형 차트로 나타내었다. 2010년~2017년까지는 2인 저자가 가장 많았으며, 3인, 1인 순으로 나타났으며, 1인 저자인 경우는 연도별 추이가 다소 줄어들어 가는 것을 볼 수 있다.

연구의 주저자와 교신저자와의 관계를 통해 몇 개의 집단으로 구성되는지 해당 집단에서 영향력이 있는 저자는 누구인지(누가 유력자인지)에 대한 관계 자료를 추출하여 시각화하였다. 그림 22와 같이 6개의 큰 그룹과 다수의 소그룹으로 이뤄져 있으며, 큰 그룹의 교신저자로서 활동만 보았을 때 탁진국(원용재,

표 1. 논문의 저자 수

	1인	2인	3인	4인	5인	8인	총
2010	6	19	5	1	2		33
2011	7	15	9	2			33
2012	3	22	8	2	2		37
2013	3	11	8	3	1		26
2014	2	17	6	5	1		31
2015	2	17	9	4		1	33
2016	1	13	5	4	2		25
2017	2	12	8	3	1		26
총	26	126	58	24	9	1	244
(%)	10.7	51.6	23.8	9.8	3.7	0.4	100

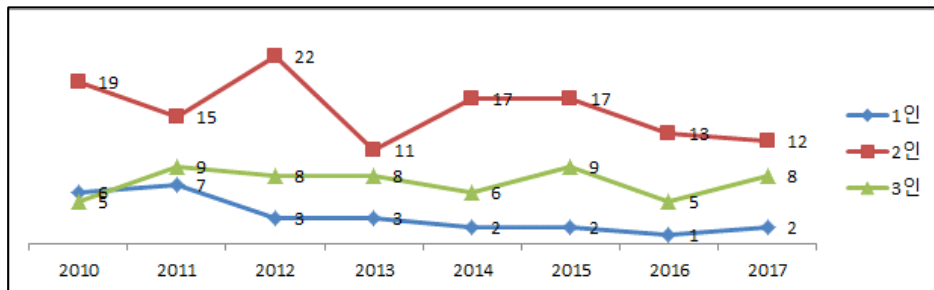


그림 21. 연도별 1인, 2인, 3인에 대한 동향

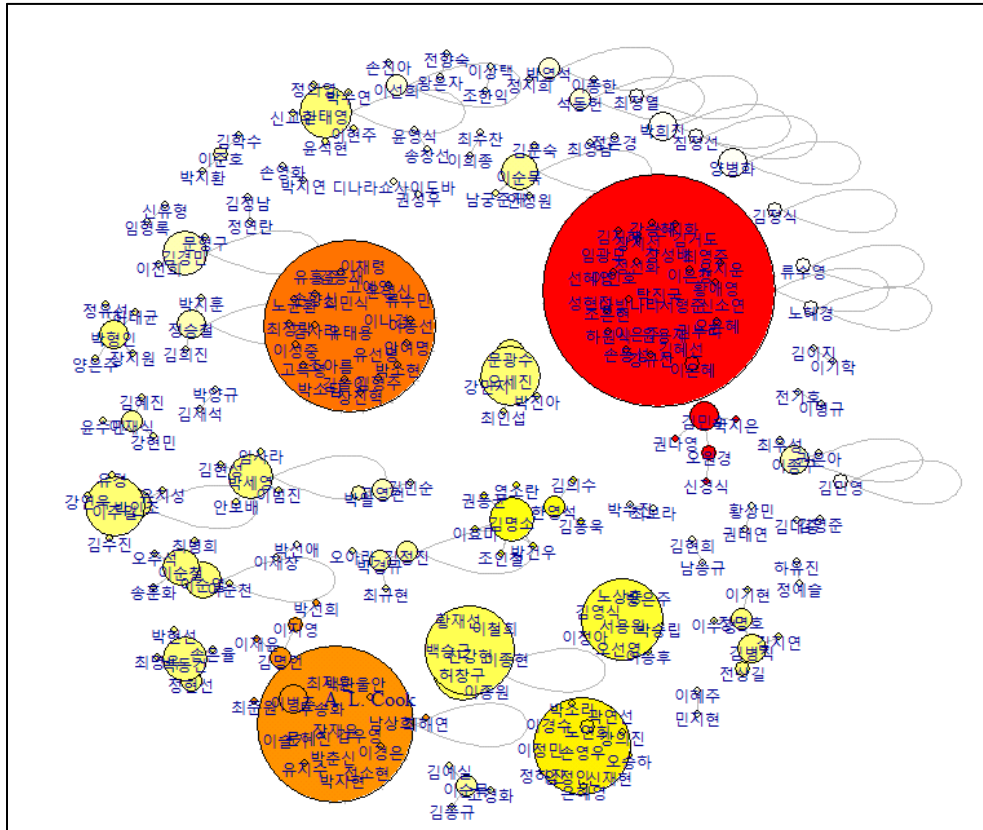


그림 22. 2010~2017년 학회지 영향력집단

탁진국, 2017; 이은경, 탁진국, 2010), 유태용 (최민식, 유태용, 2017; 안여명, 유태용, 2010), 장재윤(장재윤, 2010; 전소현, 장재윤, 이혜주, 2017), 신강현(백승근, 신강현, 김완석, 유지훈, 한영석, 2010; 이종원, 신강현, 이종현, 김완석, 2016), 서용원(김영식, 한성호, 서용원, 2017; 노상충, 정홍식, 서용원, 2012), 손영우(강의진, 정은경, 손영우, 2011; 이경수, 임정인, 박지영, 손영우, 2017) 등이 꾸준히 논문을 게재하면서 이 그룹의 유력한 조력자로서 활발한 활동을 한 것으로 나타났다. 여기서 주저자 단독으로 된 논문인 경우는 그림 22와 같이 화살표가 자기 자신을 가리키도록 하였으며, 연결된 저

자의 수가 많을수록 원의 크기가 크며 색깔이 짙게 나타나도록 나타냈다. 그림 23~그림 25는 그림 22의 2010~2017년 학회지 영향력 집단에서 교신저자로서 활발한 활동을 하고 있는 3인인 탁진국, 유태용, 장재윤 교신저자의 주 저자와의 관계를 시각화하였다. 논문의 저자는 주저자, 공동저자, 교신저자로 분류되고, 논문의 저자 표시는 일반적으로 주저자, 공동저자, 교신저자 순으로 작성한다. 분석한 논문의 79%는 교신저자를 마지막에 기재하였지만, 21%는 규칙적으로 기재되지 않았다. 그림 26은 논문에서 교신저자의 위치를 마지막에 기재하지 않은 경우를 연도별로 나타냈다. 2016

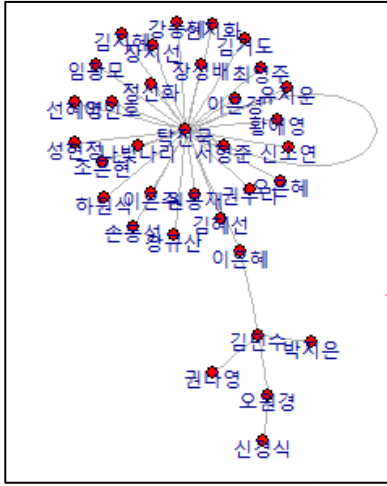


그림 23. 탁진국 교신저자 영향력

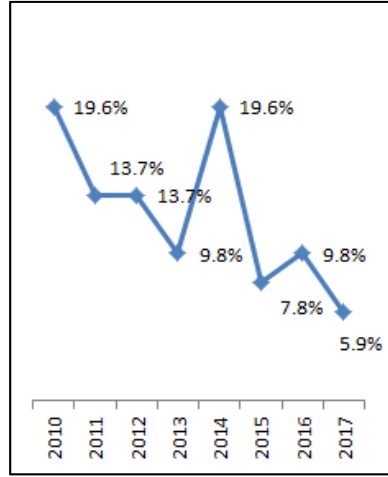


그림 26. 교신저자의 표기위치

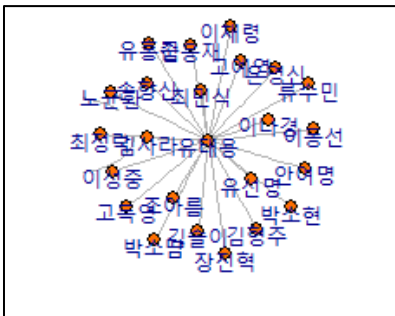


그림 24. 유태웅 교신저자 영향력

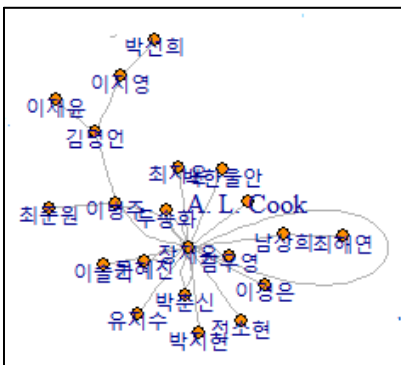


그림 25. 장재원 교신저자 영향력

년(7.8%) 2017년(5.9%)로 점차 줄어들고 있는 추세이다.

논 의

최근 빅데이터가 보급되면서 대용량 텍스트 분석기술에 관한 관심이 증가하고 있으며, 방대한 양의 학술 데이터에서 가치 있는 정보를 도출해 내는 연구들이 수행되고 있다. 본 연구는 2010년부터 2017년까지 “한국심리학회지: 산업 및 조직”에 게재된 총 244편의 논문을 대상으로 소요 시간과 노력을 줄일 수 있는 대용량 텍스트 분석기술을 이용하여 자동으로 분류하고, 분석결과를 전달하기 쉽게 시각화하는 새로운 접근방법을 시도하였다.

본 연구의 결과분석에 따라 세 가지로 결론을 정리하고자 한다. 첫째, 저자들이 논문에서 중요하다고 생각하는 용어인 주요어를 시각화한 결과 ‘조직몰입, 직무만족, 직무열의, 이직의도, 정서노동, 직무탈진, 조직시민행동’ 순으

로 두드러지게 나타났다. 전반적으로 '조직몰입과 직무만족'이 가장 관심 주요어로 나타났다. '조직몰입'은 조직 구성원이 자신이 속한 조직에 대해 가지는 심리적 애착 및 충성심을 느끼는 정도를 나타내는 것으로, 일반적으로 조직원과 조직 간의 관계를 나타낸다(Allen & Meyer, 1990). '직무만족'은 조직원이 자신의 직무를 자신이 평가하거나 직무 그 자체를 통해 얻어지는 유쾌함이나 긍정적 감정 상태를 말한다(Locke, 1976). '조직몰입과 직무만족'은 각 분야별로 비중에서는 차이가 있지만 공통된 주제이기 때문에 논문에서 저자들이 주요어로 많이 사용한 것으로 보인다. '직무열의'는 업무에 대한 긍정적이고 정력적이며 헌신하고 몰두하는 자세이다(Schaufeli & Bakker, 2004). 조직몰입과 직무만족관련 주요어에 비해서 사용빈도가 적기는 하지만 유사한 개념으로 저자들이 관심 갖는 공통된 주제로 보인다. 다음으로 '이직의도'가 관심 주요어로 나타났다. 이직은 직무수행 준거 중의 하나로 고용의 안정성을 측정하기 위한 것이다. 따라서 연구자들은 조직원이 이직을 하는 이유를 찾고자하는 목적으로 이직의도에 관심을 가진 것으로 보인다. 이직의도 다음으로 '정서노동'이 관심 주요어로 나타났다. '정서노동'이란 직장에서 조직원이 직무를 수행하면서 자신이 느끼는 본래와는 다른 정서를 의무적으로 표현해야하는 행동이다. 이것 또한 직무수행을 측정하기 위한 준거의 하나로 관심 있게 다룬 것으로 보인다. '직무탈진'이란 의욕감소, 피곤, 절망, 신체적 소진 증세 등으로 나타나는 정서적이고 신체적인 고갈상태로 나타나는 스트레스의 특수한 유형이다(Cordes & Dougherty, 1993). '조직시민행동'은 조직 구성원 스스로가 조직을 위해 행하는 자발적인 행동으로, 조직

의 공식적 보상체계에 의해 보상되지는 않으나, 전체적으로 합하여 조직이 효과적으로 기능하는데 도움이 되는 행동을 말한다(Organ, 1988). 2000년대 이후로 지금까지 일의 의미는 변하지 않았더라도 직업에 대한 관점을 다각도로 보아야 한다는 생각을 하도록 만들었다. 현 정부도 일자리 창출을 우선 공약으로 내세울 정도로 일자리 문제는 중요한 과제로 떠오르고 있다. 이에 I-O 심리학에서는 4차 산업혁명으로 인해 일자리에 대한 위기가 아닌 일자리 창출의 새로운 기회가 될 수 있는 방법을 찾고 다양한 각도로 주제를 다루는 것이 필요할 것이다. 조직원들은 자신의 일터에서 경쟁력을 유지하기 위해서는 변화를 기꺼이 수용하고 변화를 효과적으로 다룰 수 있는 능력을 지니고 있어야 한다. 이러한 변화는 I-O 심리학이 다루는 내용 자체에 영향을 미치고 있다.

2010년부터 2017년까지 연도별로 산업 및 조직심리학회의 주요어 변화를 보면 다음과 같다. 시각화 결과 '조직몰입'이 2011년, 2013년, 2014년, 2015년의 주요어로 꾸준하게 나타났다. 이는 조직몰입의 중요성을 인식하고 연구함으로써, 조직의 효과성을 높일 뿐만 아니라 일을 하는 사람에게 만족감과 삶의 보람을 얻는데 이바지하는 것으로 보인다. 반면에 2017년도에는 '이직의도'가 관심 주요어로 나타났다. '이직의도'는 Meyer, Allen과 Smith (1993)의 조직몰입의 3요소인 정서적 요소, 지속적 요소, 규범적 요소 중 지속적 요소와 가장 관련성이 높은 것으로 나타난다. 지속적 요소는 조직원이 자기가 속한 조직을 떠나면 손해라는 생각에서 조직에서 계속 머무는 몰입을 나타낸다(Meyer, et al., 1993). 규범적 요소 또한 이직의도에 영향을 줄 수 있는데, 조직원이 조직에 있는 이유가 자신이 그 조직에

있어야 하는 것이 옳다는 조직원의 가치에서 비롯된 것으로 보인다.

2017년에 관심 주요어가 ‘이직의도’ 이외에 ‘관계갈등’, ‘과업갈등’, ‘일의 의미’ 등으로 볼 때, 관계갈등과 과업갈등으로 인해서 자기가 속한 조직에 계속 머무는 것에 대해서 회의를 느낄 수 있다. 또한 조직원의 가치관의 변화로 일이 자신에 대해서 어떤 의미인지를 고민하고, 조직에 남아야 하는 이유를 찾으면서 이직에 대한 관심을 불러일으킬 수 있다. 이것은 2016년에 ‘적응수행’, ‘직무 스트레스’로 인해 이직이 주요어로 나타나기 시작하여, 2017년 이직에 대한 관심이 부각되어 주요어에 영향을 준 것으로 보인다. 직업은 특정분야의 필수적인 기술, 지식, 활동들의 집합을 나타내며, 같은 직업을 유지하면서도 다른 조직들로 이직하는 것이 가능하다. ‘이직의도’는 다양한 각도로 해석할 수 있다. 대중매체를 통해 4차 산업혁명에 관한 정보를 자주 접하면서, 직업에 대한 불안감과 위기의식으로 인해 저자에게 이직에 대한 관심을 일깨운 것으로 보인다.

연구방법을 주요어로 사용하는 경우는 흔하지 않지만, 2014년 ‘메타분석’이 관심 주요어로 나타났다. 논문에서 매개효과와 조절효과를 이용한 분석이 대부분이었으나, 2014년은 새로운 방식이라고 할 수 있는 메타분석에 관심을 가졌음을 알 수 있다. 그러므로 연구에 대한 새로운 접근 방법을 모색해 볼 필요가 있음을 시사한다.

둘째, 논문의 ‘제목과 초록’을 분석하여 순위와 관계없이 주제 10개를 선정하였다. 본 연구에서 나타난 주제는 다음과 같다. 1번째 주제는 ‘직무, 만족, 심리’를 중심으로 ‘리더십, 긍정, 자본’에 관한 단어로 이뤄진 주제이다.

이것은 I-O 심리학의 분야 중 작업생활의 질, 조직개발 등과 관련이 있는 것으로 보인다. 2번째 주제는 ‘직무, 정서, 노동, 열의’를 중심으로 ‘탈진, 종업원, 유형’에 관한 단어로 이뤄진 주제이다. 이것은 I-O 심리학의 분야 중 조직개발, 교육과 개발 등과 관련이 있는 것으로 보인다. 3번째 주제는 ‘자기, 적응, 탐색, 구조’를 중심으로 ‘근로자, 수준, 효능, 조직, 몰입, 변화, 평가, 정체성’에 관한 단어로 이뤄진 주제이다. 이것은 I-O 심리학의 분야 중 조직개발, 직무수행평가 등과 관련이 있는 것으로 보인다. 4번째 주제는 ‘갈등, 직장, 가정, 집단, 과업’을 중심으로 ‘구성, 성과, 역할’에 관한 단어로 이뤄진 주제이다. 이것은 I-O 심리학의 분야 중 조직개발, 작업생활의 질 등과 관련이 있는 것으로 보인다. 5번째 주제는 ‘행동, 관리, 상사, 동기, 침묵, 종업원, 자기’를 중심으로 ‘인상, 창의, 부하, 지향성, 학습목표, 조직, 지원’에 관한 단어로 이뤄진 주제이다. 이것은 I-O 심리학의 분야 중 조직개발, 교육과 개발, 작업생활의 질 등과 관련이 있는 것으로 보인다. 6번째 주제는 ‘이직, 직무’를 중심으로 ‘조직, 평가, 관련, 의도, 직업, 특성, 개발’에 관한 단어로 이뤄진 주제이다. 이것은 I-O 심리학의 분야 중 직무수행평가, 조직개발, 작업생활의 질 등과 관련이 있는 것으로 보인다. 7번째 주제는 ‘행동, 성격, 리더십, 심리’를 중심으로 ‘위험, 조직, 특성, 혁신, 운전’에 관한 단어로 이뤄진 주제이다. 이것은 I-O 심리학의 분야 중 교육과 개발, 작업생활의 질, 조직개발, 인간공학 등과 관련이 있는 것으로 보인다. 8번째 주제는 ‘조직, 몰입, 행동, 지각’을 중심으로 ‘안전, 구성, 시민’에 관한 단어로 이뤄진 주제이다. 이것은 I-O 심리학의 분야 중 교육과 개발, 작업생활의 질, 조직개

발, 인간공학 등과 관련이 있는 것으로 보인다. 9번째 주제는 ‘경력, 개발, 역량, 활동, 성공’을 중심으로 ‘주관, 측정, 반응’에 관한 단어로 이뤄진 주제이다. 이것은 I-O 심리학의 분야 중 교육과 개발, 직무수행평가 등과 관련이 있는 것으로 보인다. 10번째 주제는 ‘행동, 개인, 조직’을 중심으로 ‘목표지향성, 주의, 지속, 학습, 학습활동’에 관한 단어로 이뤄진 주제이다. 이것은 I-O 심리학의 분야 중 조직개발, 작업생활의 질 등과 관련이 있는 것으로 보인다.

I-O심리학의 연구 분야는 Muchinsky(2011)의 연구 분야의 기준에서 보면 몇 가지 한정된 주제로만 연구가 진행되는 것으로 나타난다. 교육과 개발, 직무수행평가, 조직개발, 작업생활의 질 분야에 치우쳐 있는 경향이 있으며, 선발과 배치, 인간공학 분야는 미흡한 것으로 드러났다. 조직원의 선발과 배치는 조직원들이 자신에게 적합한 직업에 종사할 수 있도록 만드는 우선단계라고 할 수 있다. 그러므로 선발과 배치 분야의 연구가 필요한 것으로 보인다. 또한 인간공학은 생리학, 산업의학, 인간의 지각 등에서 얻어진 지식을 사용해야 하는 분야이기 때문에 연구자가 다양한 분야의 지식을 가지고 효과적으로 조작할 수 있는 작업체계를 설계하는 것이 쉽지는 않을 것이다. 논문의 분석결과 안전에 관한 것이 연구 내용이 포함되어 있지만 미비한 상황이다. 그러므로 작업현장에서 발생하는 다양한 사고들을 줄이기 위해서도 위험상황을 사전에 감지하고 환경을 구성할 수 있는 연구가 필요한 것으로 보인다.

미국 I-O 심리학의 트렌드 분석에서 보면 빅 데이터 분석, 기술, 일과 삶 등이 2014년부터 2017년까지 꾸준히 트렌드로 나타났다. 기

술(직장에서의 기술 통합, 업무 수행 방식에 대한 기술의 시사점 증가, 기술 동향)과 일과 삶(일과 삶의 균형, 일과 삶의 통합, 여러 세대의 일과 삶의 균형)에 관한 내용은 부분적으로 한국 I-O 심리학에서 주제로 나타났지만, 빅데이터(빅 데이터, HR 분석 및 빅 데이터의 지속적인 사용, 빅 데이터 활용 및 극대화, 빅 데이터를 활용하여 데이터 중심 의사 결정)에 대한 주제의 언급은 메타분석을 제외하고는 중요 주제로 나타나지 않았다. 그러므로 “한국 심리학회지: 산업 및 조직”의 주요어와 미국 I-O 심리학의 주요 트렌드를 비교하여 볼 때 빅데이터를 활용한 연구들에 대한 관심을 가질 필요가 있다.

산업 및 조직 심리학자들은 심리적으로 안전하고 건강한 근무 환경이 되도록 변화하는 작업장의 동향을 파악하고 기업이 해결하고자 하는 것을 예측하는 것이 필요하다. 이에 역동적으로 변화하는 사회에서 구성원들에게 변화를 인식할 수 있도록 산업 및 조직 심리학자들은 트렌드를 분석하여 제시하는 것도 하나의 방법이다. 그러나 본 연구 결과, 전반적으로 국내 산업 및 조직 심리학자들은 전통 I-O 심리학 분야에 국한된 주제로 연구하는 것으로 나타났다. I-O 심리학은 4차 산업혁명 시대에 가장 민감하게 대응해야함에도 불구하고 변화하는 상황을 반영한 연구의 시도가 미비한 것으로 보인다. 반면에 2014년 미국 산업 및 조직심리학회 10대 트렌드 중 ‘게이미피케이션(Gamification)’을 선정하였다. 게이미피케이션은 게임이 아닌 분야에 대한 지식 전달, 행동 및 관심 유도 혹은 마케팅 등에 게임의 매커니즘, 사고방식과 같은 게임의 요소를 접목시킬 수 있다는 것이다(Jane, 2011). 이것은 새로운 패러다임이라고도 할 수 있을 만

크의 작업 공간과 작업 변화의 가능성을 보여 준 것이라 할 수 있겠다.

셋째, 저자의 수는 2인이 가장 많았으며, 3인, 1인 순으로 나타났고, 1인 저자의 논문은 감소하는 추세이다. 유태용과 탁진국(1998)의 연구 결과에 따르면, 1997년 이전에는 전반적으로 공동연구보다는 단독연구의 비중이 컸지만, 본 연구 결과에서는 2~3인으로 이루어진 공동연구가 더 비중이 큰 것으로 나타났다. 창의적 기초능력을 배양하자는 시대적 요구에 따라 단독연구보다는 공동연구를 더 많이 하는 것으로 보인다. 그리고 연구의 주저자와 교신저자와의 관계를 시각화하였다. 6개의 큰 그룹과 다수의 소그룹으로 이뤄져 있으며, 큰 그룹의 교신저자로서 ‘탁진국, 유태용, 장재윤, 신강현, 서용원, 손영우’ 등이 그룹의 유력한 조력자로서 활발한 활동을 한 것으로 나타났다. 유력한 조력자들이 관심 갖는 연구주제는 I-O 심리학의 분야를 선도하는 주제라고 할 수는 없지만 I-O 심리학의 발달적 측면에서는 가치가 있다고 할 수 있을 것이다. 2010년부터 게재된 논문의 수는 25~37편으로 연평균 31편이었다. 미국 노동부 통계에 따르면 2012년부터 20년간 미국에서 I-O 심리학이 가장 빠르게 성장하는 분야의 1위를 차지할 정도지만 본 연구결과 2016년, 2017년 학회지의 게재된 논문 편수는 평균보다 저조한 양상을 보이는 것으로 보아 연구 활동이 활발했다고 보기는 어렵다.

본 연구의 한계점은 다음과 같다. 첫째, 대용량 텍스트 분석기술을 이용하여 쉽고 빠르게 자동으로 주제를 분류할 수 있어서 시간과 노력을 감소시켜 변화하는 기술에 민첩하게 대응할 수 있다는 장점은 있지만, 평정자가 직접 평정하여 분류·분석하는 기존의 방식에

비하여 정확도는 다소 낮을 수도 있을 것이다.

둘째, 본 연구에서 실시한 방법은 비정량 자료를 시각화하여 알기 쉽게 전달되도록 하는 것을 목적으로 하였기 때문에 결과의 통찰과 활용 방안은 관련 연구자의 역량에 좌우된다.

셋째, 논문에서 교신저자의 표기 위치가 21%(51편)이 일정하지 않고, 불규칙적으로 기재되어 있어서 확인 후 일일이 처리하는 작업이 이뤄졌다. 최근 2017년도에는 5.9%로 교신저자의 표기 위치를 어느 정도의 원칙을 정한 것으로 보이나 학회의 논문 양식에서 교신저자의 위치를 명시하는 것이 좋을 듯하다. 본 연구는 시간과 노력을 최소화하는 것을 목적으로 자동적으로 주제 분류를 진행하였으나, 주저자와 교신저자의 관계는 표기 위치로 인해 수작업을 병행할 수밖에 없었다.

참고문헌

- 강의진, 정은경, 손영우 (2011). 부하 직원의 동기가 수행에 미치는 영향-상사의 감정 표현에 대한 귀인 방식의 조절효과를 중심으로. 한국심리학회지: 산업 및 조직, 24(3), 553-574.
- 김영식, 한성호, 서용원 (2017). 과업갈등이 관계갈등에 미치는 영향에 대한 조절효과-관계의 질과 개인차에 따라서. 한국심리학회지: 산업 및 조직, 30(3), 415-441.
- 노상충, 정홍식, 서용원 (2012). 글로벌 역량 척도(GCS) 개발 및 타당화. 한국심리학회지: 산업 및 조직, 25(4), 801-831.
- 백승근, 신강현, 김완석, 유지훈, 한영석 (2010). 조직의 정서표현규칙이 종업원의 직무태도에 미치는 영향 - 정서노동 전략의 매

- 개효과를 중심으로, 한국심리학회지: 산업 및 조직, 23(1), 27-51.
- 안여명, 유태용 (2010). 개인 및 팀 수준에서 성격과 적응수행 간의 관계 - 일반적 자기효능감의 매개효과와 변혁적 리더십의 조절효과. 한국심리학회지: 산업 및 조직, 23(1), 155-179.
- 원용재, 탁진국 (2017). 조직변화가 직무스트레스에 미치는 영향에서 고용불안의 매개효과: 무경계경력태도의 조절효과를 중심으로, 한국심리학회지: 산업 및 조직 30(4), 563-587.
- 유태용, 탁진국 (1998). 창간 10주년을 맞는 '한국심리학회지: 산업 및 조직'의 내용분석 및 앞으로의 제언, 한국심리학회지: 산업 및 조직, 11(1), 53-76.
- 이경수, 임정인, 박지영, 손영우 (2017). 공군 전투비행팀의 과업갈등, 관계갈등, 팀효능감 및 과업성과의 관계 - 참여적 의사결정의 조절효과, 한국심리학회지: 산업 및 조직, 30(3), 329-355.
- 이은경, 탁진국 (2010). 상사의 정서지능이 부하의 조직몰입, 직무만족에 미치는 영향 - 리더행동의 매개효과를 중심으로, 한국심리학회지: 산업 및 조직, 23(1), 1-25.
- 이종원, 신강현, 이종현, 김완석 (2016). 조직정치 지각에 대한 개인-조직 부합도와 마키아벨리즘의 효과, 한국심리학회지: 산업 및 조직, 29(3), 385-410.
- 잠재 디리클레 할당, 위키피디아, 2017.1.5. https://ko.wikipedia.org/wiki/%EC%9E%A0%EC%9E%AC_%EB%94%94%EB%A6%AC%ED%81%B4%EB%A0%88_%ED%95%A0%EB%8B%B9에서 자료 얻음.
- 장재윤 (2010). 직무만족과 직무에서 경험하는 정서간의 관계 - 빈도 가설과 긍정성 비율 가설의 비교, 한국심리학회지: 산업 및 조직, 23(2), 275-295.
- 전소현, 장재윤, 이해주 (2017). 공격성에 대한 조건추론검사(CRT-A)의 예측타당도연구 - 외현적 공격성과의 상호작용 효과, 한국심리학회지: 산업 및 조직, 30(4), 541-562.
- 정재삼, 강민진, 이해진, 조지현 (2014). 학회지 분석을 통해 본 HRD 이슈와 동향 탐색 - 기업교육연구, 인적자원관리연구, 한국심리학회지: 산업 및 조직을 중심으로, 한국 기업교육학회 학술대회, 2, 37-53.
- 최민식, 유태용 (2017). 완벽주의가 능동-수동 지연행동에 미치는 영향에서 목표지향성의 매개효과, 한국심리학회지: 산업 및 조직, 30(2), 195-219.
- 한국정보화진흥원 (2012). 성공적인 빅데이터 활용을 위한 3대 요소: 자원, 기술, 인력, IT & Future Strategy 제 3호.
- Allen, N. J. & Meyer, J. P. (1990). The measurement and antecedents of affective, continuance and normative commitment to the organisation. *Journal of Occupational Psychology*, 63(1), 1-18.
- Blei, D. M. (2012). *Probabilistic topic models. Communications of the ACM*, 55(4), 77-84.
- Blei, D. M., Ng. A. Y., Jordan. M., I. (2003). Latent dirichlet allocation, *Journal of Machine Learning Research* 3, 993-1022.
- Bureau of Labor Statistics (2013. 12. 19). Bureau of Labor Statistics News release, https://www.bls.gov/news.release/archives/ecopro_12192013.pdf에서 2017. 12. 3자료 얻음.
- Cordes, C. L. & Dougherty, T. D. (1993). A review and an integration of research on job

- burnout. *Academy of Management Review*, 18, 621-656.
- Jane McGonigal (2011). Make it GAMEFUL. *The Gamification Summit-Conference Sessions*.
- Kao, A., and Poteet, S. (2007). *Natural Language Processing and Text Mining*. London: Springer-Verlag.
- Locke, E. A. (1976). The Nature and Causes of Job Satisfaction. In Dunnette, M.D.(Ed.), *Handbook of Industrial and Organizational Psychology*, Chicago: Rand McNally, 1297-1343.
- Meyer, J. P., Allen, N. J., Smith, C. A. (1993). Commitment to organizations and occupations: Extension and test of a three-component conceptualization. *Journal of Applied Psychology*, 78(4), 538-551.
- Muchinsky, P. M. (2012). 산업 및 조직심리학 [Psychology Applied to Work], (유태용 역), 서울: 시그마프레스(원전은 2011에 출판).
- Organ, D. W. (1988). *Organizational Citizenship Behavior - The Good Soldier Syndrome*, MA: Lexington Books.
- Schaufeli, W. B., & Bakker, A. B. (2004). Job demands, job resources and their relationship with burnout and engagement: A multi-study, *Journal of Organizational Behavior*, 25(3), 293-315.
- Society for Industrial and Organizational Psychology (2014). Announce the Top 10 Workplace Trends for 2014, *EHS Today*, 7(3), 12.
- Canadian Psychological Association (2015). Top 10 workplace trends for 2015, *Canada's Psychology Newspaper*, 37(4), 18-19.
- Society for Industrial and Organizational Psychology (2016). Top 10 workplace trends for 2016, *EHS Today*, 9(1), 10.
- Society for Industrial and Organizational Psychology (2017). SIOP Announces Top 10 Workplace Trends for 2017, *The Industrial -Organizational Psychologist*, 54(3), 139-143.
- 1차 원고접수 : 2018. 01. 26
2차 원고접수 : 2018. 05. 11
최종게재결정 : 2018. 05. 11

Unstructured Data Analysis and Visualization “Korean Journal of Industrial and Organizational Psychology(2010~2017)”

Insun Yi

Yeungnam Uni.

Eunyoung Na

Andong Uni.

With the spread of Big data, there is an increasing interest in Unstructured Data analysis techniques, and studies are being carried out to extract valuable information from vast amounts of academic data. Most passive analysis of large texts by humans is time consuming and laborious and difficult, so a technique that can automatically classify them in order to supplement them is needed. Therefore, this study analyzed and visualized by using Unstructured Data such as the main words, thesis title, and abstract of ‘Korean Journal of Industrial and Organizational Psychology’ published from 2010 to 2017. The results of the study are as follows. First, ‘organizational commitment’ was frequently used as a result of visualization of the key words using wordcloud. Job satisfaction, job enthusiasm, turnover intention, emotional labor showed the order. Second, the title of the paper and the abstract were automatically classified into 10 topics based on the LDA probability. Topic 8(organization/commitment/perception) was the highest, and Topic 5(behavior/management/boss) was the lowest. Third, the relationship between the main author and the correspondent author appeared as six large groups and several small groups. We were able to identify influential authors within the group. In this study, it is suggested that related researchers can get access to another research by deriving information from a vast amount of academic data more quickly and easily.

Key words : Unstructured Data, LDA, Bing Data, I-O Psychology