

머신 데이터 분석용 플랫폼 스플링크를 이용한 취업지원 서비스 개선에 관한 연구 : 월드잡플러스 사례를 중심으로

이재덕¹, 이문기², 김미량^{3*}

¹한국산업인력공단 정보화지원국, ²성균관대학교 경영대학, ³성균관대학교 컴퓨터교육과

Experiencing with Splunk, a Platform for Analyzing Machine Data, for Improving Recruitment Support Services in WorldJob+

Jae Deug Lee¹, MoonKi Kyle Rhee², Mi Ryang Kim^{3*}

¹Information Support Bureau, Human Resources Development Service of Korea

²School of Business, SungKyunKwan University

³Dept. of Computer Education, SungKyunKwan University

요 약 한국산업인력공단이 운영하는 월드잡플러스는 청년들의 해외취업을 지원하는 포털 서비스로서 해외진출에 필요한 정보제공과 등록, 면접, 학습 등 일련의 과정을 지원하는 통합정보네트워크이다. 현재 30만명 이상의 청년들이 등록하고 있으며, 연계관련기관과 협업하여 청년들의 해외 취업을 지원한다. 월드잡플러스는 지원서비스의 혁신화와 무결성 유지를 위해 머신데이터 분석플랫폼인 스플링크를 활용하여 웹사이트에 축적된 로그파일 분석을 시도하고 있다. 기술적·예측적 분석 도구를 이용하여 구직자 니즈와 프로필 기반의 맞춤형 매칭 서비스를 제공하며 구직자를 위한 최적 구직 성공요건 및 최적 구인기업에 대한 정보를 제공한다. 본 논문에서는 월드잡플러스가 스플링크를 활용하여 해외취업을 지원하는 몇 가지 서비스에 대한 사례를 제시해보고자 한다.

주제어 : 빅데이터 분석도구, 월드잡플러스, 스플링크, 취업지원서비스, 시각화

Abstract WorldJob+, being operated by The Human Resources Development Service of Korea, provides a recruitment support services to overseas companies wanting to hire talented Korean applicants and interns, and support the entire course from overseas advancement information check to enrollment, interview, and learning for young job-seekers. More than 300,000 young people have registered in WorldJob+, an overseas united information network, for job placement. To innovate WorldJob+'s services for young job-seekers, Splunk, a powerful platform for analyzing machine data, was introduced to collate and view system log files collected from its website. Leveraging Splunk's built-in data visualization and analytical features, WorldJob+ has built custom tools to gain insight into the operation of the recruitment supporting service system and to increase its integrity. Use cases include descriptive and predictive analytics for matching up services to allow employers and job seekers to be matched based on their respective needs and profiles, and connect jobseekers with the best recruiters and employers on the market, helping job seekers secure the best jobs fast. This paper will cover the numerous ways WorldJob+ has leveraged Splunk to improve its recruitment supporting services.

Key Words : Big-data analytics, WorldJob+, Splunk, Recruitment service, Visualization

*Corresponding Author : Mi Ryang Kim(mrkim@skku.ac.kr)

Received January 10, 2018

Accepted March 20, 2018

Revised February 23, 2018

Published March 28, 2018

1. 서론

엄청난 속도로 발전하는 인터넷 기술과 다양한 소셜 미디어의 등장은 데이터가 경제적 자산이 되는 빅데이터 시대의 도래를 가능하게 하였다. 컴퓨터 스토리지와 CPU관련 기술이 뒷받침하지 못했던 2,000년대 초반에도 데이터의 양산은 많은 기업과 조직에게 큰 부담으로 작용했다. 분석기법을 뒷받침할만한 컴퓨터 기술도 충분하지 못했고 분석방법 역시 결과를 효율적으로 활용할만큼 개발이 되지 못했기 때문이다. 그러나 이제 단순히 정보를 주고받던 웹1.0에서 집단 지성을 창출하고자 했던 웹 2.0을 지나 데이터분석을 통해 영화나 소셜에서나 상상할 수 있었던 융합지식 창출을 가능하게 하는 웹 3.0시대가 등장하였다.

IT관련 연구기관인 가트너(Gartner)에 의하면 빅데이터의 특징은 단순히 데이터가 엄청나게 크다는 데에 있지 않다는 것을 강조한 바 있다. 빅데이터는 양뿐만 아니라 다양성과 변화의 속도로 인해 그 가치가 인정받는다 [1]. 오늘날 인터넷을 통해 쏟아져 나오는 데이터의 형태가 매우 다양하기 때문에 단순히 큰 규모로만 빅데이터라고 할 수 없으며 전혀 정형화되지 않은 SNS상의 댓글, 뉴스나 커뮤니티 사이트의 게시물, 동영상, 팟캐스트, 음악, 사진 등까지 포함한다. 이처럼 빅데이터는 비정형화 혹은 비구조적 상태일 뿐만 아니라 데이터가 쌓이는 속도와 내용의 변화속도 역시 가능하기 힘들다.

빅데이터의 가치는 계량적인 분석도구를 통해 다양하면서 큰 규모의 데이터를 활용해 이전에는 불가능했던 새로운 통찰이나 새로운 형태의 가치를 추출해 내는 데에 있다. 즉, 이전에는 볼 수 없었던 세상을 보게끔 해주고 복잡한 세상을 더욱 잘 이해할 수 있도록 도와주는 역할을 하게 될 가능성이 높다. 빅데이터의 분석과 활용은 환경에 대한 새로운 차원의 분석과 인식을 가능하게 해 줌으로써 신속하고 더 나은 의사결정으로 이끌어준다. 즉, 다양한 데이터들을 매우 빠르게 분석할 수 있게 됨으로써 다양한 정보를 지체 없이 분석할 수 있게 되어 의사결정의 질이 한 차원 제고되며 이를 통해 조직의 운영과 더욱 효율적으로 이끌어 나갈 수 있는 방법을 제공해 준다 [2].

정부를 포함하는 공공부문도 빅데이터를 기반으로 하는 융합지식 창출을 통해 정부 운영의 효율화와 투명성 제고에 많은 노력을 기울이고 있으며, 특히 맞춤형 국민

서비스 요구에 적극적으로 대응하기 위한 정책적 대안으로 빅데이터 분석을 적극 활용하고자 기획하고 있다. 예를 들면, 국내·외 경제·사회 정세를 실시간으로 파악하는 분석력을 확보하여, 질병·금융위기 등 위협에 대한 대응력을 강화하며, 데이터의 공개·활용을 통해 인건비 절감, 세수 투명성 확보, 국민복지 향상을 추구하고 있다 [3,4].

공공부문의 빅데이터에는 주민등록 데이터, 소득 관련 데이터, 의료 서비스 관련 데이터, 차량 보유 및 운행 관련 데이터 등과 같은 정형적 데이터뿐만 아니라 공공부문의 웹서비스 활용과정에서 축적되는 비정형적 빅데이터도 포함된다 [5]. 그러나 최근까지도 이런 데이터는 간단한 기술통계 정도의 분석만이 가능한 통계 데이터로서 그 존재성이 인정되었고 많은 경우 행정 프로세스의 부산물로 존재하여 데이터가 갖고 있는 경제적, 정책적 가치를 인지하지 못한채 방치되고 있었다고 평가할 수 있다 [6]. 이런 다양한 분야에서 구축되는 빅데이터의 분석은 공공부문의 혁신뿐만 아니라 개별화된 대국민 서비스도 개발할 수 있는 기반을 제공한다.

빅데이터의 활용은 글로벌기업들뿐만 아니라 효율적인 정부 운영을 도모하는 선진국들도 이 분야에 공격적으로 투자하고 있다. 정부를 포함하는 공공부문은 관료제적인 속성으로 인해 데이터의 공유가 쉽지 않다. 그러나 빅데이터 시대의 도래는 기존의 공공데이터의 수직적·단절적 활용에서 기관간 연동을 통한 수평적 활용프로세스를 혁신화함으로써 실시간 이슈를 탐지하고 선대응하는 정부를 구현하는데 상당한 역할을 한다.

미국 국립의학도서관의 필박스(Pillbox) 프로젝트가 좋은 예이다. 필박스 서비스는 미국 국립보건원에 접수되는 알약의 기능이나 유효기간을 문의하는 민원수가 너무 많아 사이트를 통해 자세하게 소개하는 서비스로서, 단순히 사용자에게 정보를 제공하는 것에 그치지 않고 사용자가 검색한 약의 정보를 통해 의약품 사용과 관련된 지도, 그래프를 입체적으로 작성해 제공한다. 오바마 정부는 이 필박스 프로젝트를 통해 연간 오천만 달러의 비용 절감 효과를 거두었다고 한다 [7,8]. 독일 연방 노동기관도 빅데이터 활용 맞춤형 고용서비스를 통해 3년간 백억 유로의 비용을 절감할 수 있었다고 하며, 우리나라도 공공 분야 빅데이터 활용시 경제 효과가 10조원 이상에 달할 것으로 예상한다 하고 있다 [4].

그렇지만 우리나라의 경우 공공부문의 빅데이터 분석 사례는 거의 찾아보기 힘들다. NIA에서 발간한 공공데이

터의 활용사례집[9]도 공공데이터를 재구조화하여 활용한 사례를 소개한 것으로 본격적인 빅데이터 분석과는 거리가 있다. 본 연구는 공공부문 빅데이터 분석의 정책적 중요성을 재차 인식시키고 일종의 예시사례로 글로벌 취업시장에서 빅데이터 분석결과의 활용방안을 소개하는데 목적이 있다. 이를 위해 한국산업인력공단이 운영하는 월드잡플러스에 축적된 청년들의 해외취업과 관련하여 축적된 데이터를 활용하는 프로세스를 제시해 보고자 한다. 빅데이터 분석을 토대로 하여 취업정보에 대한 지원자의 성별, 연령, 학력 등 구인정보를 제공하여 취업 가능성을 예측할 수 있는 서비스 분석사례를 살펴보고, 이력정보, 공고열람정보와 해외취업공고 등 키워드 기반 빅데이터를 토대로 구직자와 구인공고 정보간 관심 키워드 가중치를 활용한 매칭업(matching up) 서비스에 대한 사례도 제시하고자 한다.

공공부문에서의 빅데이터분석이 가져다 줄 수 있는 공익적 가치에 대해서는 많은 논의가 이루어지고 있지만 실질적으로 정부나 공공기관이 빅데이터 분석을 주도적으로 이끌어 가시적인 성과를 낸 사례는 찾아보기 힘들다. 본 연구는 해외취업을 목표로 하는 청년들에 대한 지원사례로서 다른 공공부문에 활용될 수 있는 여지는 많지 않지만 빅데이터 분석이 가져다 줄 수 있는 정보의 다양성과 가치를 인식시킬 수 있다는 측면에서 상당한 기여를 할 수 있을 것으로 기대한다.

2. 빅데이터 분석의 의의와 공공데이터

2.1 빅데이터 분석의 개념

빅데이터 분석기법(big data analytics)은 일반적인 통계 및 최적화기법과 사용목적은 유사하나 빅데이터를 입력자료로 한다는 점이 다르다. 빅데이터 분석기법은 business intelligence(BI)을 이끌어가는 도구로서 새로운 변화가 무엇이며 이에 대응하는 전략이 무엇인지 파악하는데 도움을 준다. 공공부문의 빅데이터분석이 기업부문의 빅데이터분석인 비즈니스 애널리틱스(business analytics)와는 목적이 다르겠지만 디지털화된 빅데이터의 분석을 통해 의사결정을 지원하는데 활용된다는 점에서는 유사하다.

빅데이터 분석기법은 분석목적에 기준으로 기술적 분석기법(descriptive Analytics), 예측적(predictive) 그리

고 처방적(prescriptive) 분석기법으로 나눌 수 있다[10]. 기술적 분석기법은 사용자들의 서비스 활용과정에서 자동적으로 창출되는 데이터를 활용한다. 서비스 관련하여 어떤 일들이 발생하고 있는지, 계획대로 진행되고 있는지, 문제가 발생했다면 어떤 유형의 문제가 왜 발생했는가 등에 대한 답을 추구한다. 이 과정을 효율화하기 위해 다양한 대시보드(dashboard)와 보고서, 그리고 조회기능(queries)을 활용하여 레포트의 창출 및 배포, 데이터의 비주얼화 기능을 제공한다. 그렇지만 기술적 분석기법은 담당자의 주관적인 리비유를 필요로 하며, 미래에 대한 예측이나 이를 극복하기 위한 방안은 제공하지 못한다.

일정 규모 이상의 데이터가 축적되면 기술적 분석과 함께 예측적 분석이 가능해진다. 예측적 분석은 용어 그대로 미래에 대한 예측을 목표로 하며, 데이터 마이닝 등을 통해 어떤 요인들이 서로 상관관계가 높은지, 패턴 인식을 통해 프로세스를 목표방향으로 수정하기 위해서는 어떤 조치를 취해야 하나 등에 관한 분석이 이루어진다. 시뮬레이션이나 복잡한 수리적 모형을 구축하여 특정 사건이 발생한다면 어떤 결과가 나타날 것인가 등에 대한 논의가 이루어지기도 한다. 어떤 기법을 활용하더라도 과거 데이터를 분석하여 패턴이나 인과관계를 알아보고 이를 통해 미래 상태를 예측해보는데 목적이 있다[11,12].

처방적 분석은 경영과학에서 흔히 등장하는 최적화(optimization) 개념과 일맥상통한다. 조직내의 의사결정은 지나치게 많은 대안 중에서 최적의 것을 선택해야 하는 경우가 많다. 이전에는 컴퓨터용량이나 알고리즘의 계산능력 미비로 인해 최적해를 구하는 것이 거의 불가능한 문제도 이제는 단 몇 시간만에 최적해를 구할 수 있는 경우가 많아졌다. 처방적 분석기법은 이처럼 수학적 최적화에 기반을 두며 주어진 제약조건하에서의 최적대안을 찾는 것이 기본적인 목표이다. 대안이 무수히 많은 경우 최적해를 찾기 위한 수학적 해법(algorithms)을 개발해야 하며 머신러닝, 통계적 기법과 연계되어 사용되기도 한다. 처방적 분석은 향후 인간 대신 자동화한 의사결정을 가능케하는 ‘자동화 분석’(automated analytics)의 등장도 촉진시킬 것이다. 예를 들어 자동화 분석 시스템은 마케팅 담당자의 결정을 거치지 않은 채 응용 프로그램을 사용하여 고객에게 최적의 ‘판촉 이메일’을 자동으로 보내도록 할 것 있으며, 구매담당자의 결정없이도 협력업체의 컴퓨터에 구매요청 데이터를 입력하게 될 것이다.

2.2 공공부문의 빅데이터 분석 관련 연구

정부를 포함한 공공분야에서 생산된 데이터는 데이터 생성기관이 특정산업영역에서 독과점적인 위치에 있음으로 인해 분석의 기초가 되는 데이터를 모집단과 최대한 유사하게 가져갈 수 있다는 성격을 지닌다[13]. 즉 평가와 예측과정에서의 오류가 최소화되며, 더불어 데이터의 속성이 어느 한쪽으로 편향되어 있을 가능성도 낮다는 특징을 지닌다. 편향된 데이터를 사용한다면 그 결과 역시 편향될 수밖에 없다는 점을 감안한다면 공공데이터는 분석결과의 신뢰성을 극대화할 수 있는 태생적 기반을 제공한다.

이런 잠재적 가치가 충분함에도 불구하고 공공분야의 빅데이터 분석은 시작단계에 불과하다. 컴퓨터 공학분야에서는 데이터 구조나 알고리즘 관련하여 다양한 연구가 진행되고 있지만, 응용과학 분야인 경영이나 행정, 복지 등의 학문분야에서는 빅데이터의 실질적인 적용이나 이를 통한 경제적 성과에 대한 연구논문은 찾아 보기 힘들다. 관련분야의 국내 연구가 걸음마 단계에 불과한데, 예를 들어 이만재[6]는 차세대 전자정부 구현과 관련하여 링크드 데이터 방식을 이용하여 부처별 공공 데이터를 통합하는데 필요한 마이닝 기법과 시각화 도구의 필요성을 주장한 바 있으며, 함유근[14]은 공공기관의 빅데이터 통합사례를 통해 비즈니스 상황에 맞는 데이터 통합방법의 선택 전략을 제시한 바 있다. 김선홍 등[15]은 사례분석으로 교통정보 데이터의 활용방안에 대해 논의한 바 있다. 이 연구에서 교통정보분석을 위한 도구로서 R을 이용하였으며, 실험을 통해 점유율과 교통량 간의 인과관계를 분석하였다. 그 외 이병엽[16] 등은 기존의 데이터 마이닝 분석 기법을 통한 소셜 미디어의 분석 형태에 대한 활용 및 분석방안을 제시한 바 있다.

정부와 공공기관이 주도하는 빅데이터 활용분야에서도 구체적인 분석 및 활용결과는 기관 내부에 한정하여 사용할 뿐, 일반인을 위한 사례연구 형식으로 발표하지는 않는 경향이 있다. 외국의 경우 흥미로운 사례들이 발표되고 있는데, 그중 가장 유명한 사례가 미국의 필박스 프로젝트이다. 필박스 사이트는 알수 없는 알약의 이름과 효과나 부작용을 알아보는데 활용된다. 알약의 겉모양과 관련 정보를 결합함으로써 사용자 시각적으로 해당 의약품을 검색하고 일회 복용량 등을 알아볼 수 있다. 일단 그 약이 어떤 약인지 확인된 후에는 브랜드 및 복제약 이름, 약성분, 미국 약품파일 인식번호 등의 정보가 제공

되며, 미국 NIH나 FDA가 보유하고 있는 의약품 정보를 검색할 수 있는 포털 사이트를 연결해주는 링크도 제공한다. 국민들은 다양한 약에 대한 정보를 손쉽게 얻을 수 있으며, 제약회사는 자사가 개발한 새로운 약에 대한 정보를 직접 제공할 수 있어 약 관련 민원처리비용을 줄이는데 기여한다. 정부도 검색서비스를 통해 얻어진 사용자 질병에 대한 통계 데이터를 활용할 수 있다는 점에서 정책적인 활용도 가능하다.

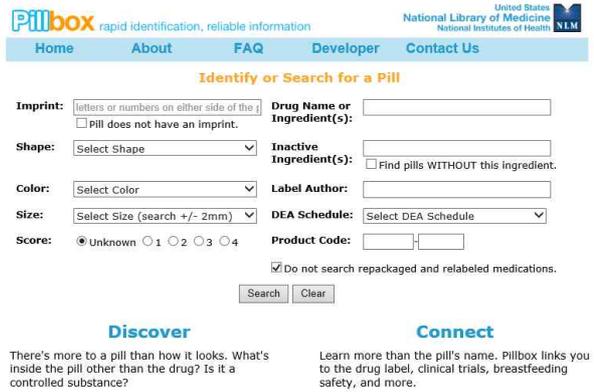


Fig. 1. PillBox Portal Site

범죄나 의학분야외에도 빅데이터분석이 가장 활발하게 활용되는 분야는 취업관련 사이트이다. 구인기업은 원하는 조건에 맞는 인재정보 검색 서비스를 제공받고, 구직자는 공공과 민간에 흩어져 있는 일자리 정보를 한 곳에서 제공받을 수 있을 뿐만 아니라 구직자 입장에서 상세한 정보를 입력할 수밖에 없어 데이터의 시의성이나 신뢰도도 매우 높은 편이다.



Fig. 2. JOBBORSE Portal Site

Fig. 2에 예시된 독일의 포탈 JOBBORSE에서는 메인 화면에서 일자리를 찾는 구직자와 직업훈련 희망자를 함께 안내한다. 이 사이트는 하나의 공통 플랫폼을 통해, 공공 취업사이트와 민간 취업사이트로부터 구인 및 구직정보를 수집함으로써 호환성을 높이고 있다. 신뢰성이 검증된 기업은 별도의 인증 절차를 생략하여 신속하게 정보를 등록하도록 유도함으로써 효율성을 높이고 데이터 분석을 통해 구직자와 구인기업 간의 매치 성공률을 높이는 서비스를 제공한다.

3. 빅데이터 분석 사례의 배경

3.1 데이터 창출기관의 소개

월드잡플러스는 해외취업·인턴·연수·봉사·창업 등 청년들의 해외진출정보를 통합으로 제공하는 포털로서 해외진출에 필요한 원스톱 맞춤형 기능과 무역협회, KOTRA, KOICA 등 관련기관과 연계하여 청년들의 해외 취업을 지원한다. 2008년 9월 글로벌리더 종합정보망 구축 사업을 통해 글로벌사업수행 모니터링 및 통계 관리를 시작한 이래 2013년 해외진출 우수 사례, 인턴, 봉사, 취업 등 해외진출 관련 정보를 제공하기 시작하였으며, 2014년에는 KOTRA, KOICA, 교육부 등의 유관기관과 연계하여 데이터 표준화 작업을 시작하였다. 2015년에는 정부해외인턴 사업을 통합하고 모바일 앱도 구축하여 운영하고 있다[17].

Fig. 3의 월드잡플러스 웹사이트를 통해 축적된 데이터에는 해외취업·연수·인턴 등 월드잡플러스 서비스에서 발생하는 방문정보, 회원가입정보, 로그인 정보, 이력서 정보, 공고 및 조회 정보, 취업지원 정보 합격, 상담내역의 기본 정보 등이 포함된다. 2017년 현재 회원수는 47만명에 이르고 있으며 해외취업 개설공고수는 2017년 7천건(누적 3만건)을 넘어섰다. 지원자수는 2017년 10만명(누적 60만명)이 넘었으며 취업자수는 2017년 3,435명으로 취업률은 3.4%를 약간 상회하는 정도이다. 해외연수는 2017년 약200건 공고에 만명 이상이 지원하여 7,070명이 합격했으며 이중 취업으로 이어진 사례는 2017년 12월 초 기준으로 2,345명으로 약 33.1%의 취업률을 보였다[17].

이력서, 공고열람정보, 지원이력, 검색어 분석 등 구직자가 사용한 데이터에 상관관계 파악 알고리즘을 활용하면 맞춤형 해외정보 서비스 제공이 가능해진다. 특히 구

직자 관련 핵심키워드를 추출 및 분석함으로써 구직활동의 특성을 파악하고 개인의 선호도를 반영할 수 있는 맞춤형 구직공고 제시를 할 수 있다.



Fig. 3. WorldJob+ Website

3.2 사례분석에 활용된 분석도구

본 연구에서 사용한 분석도구는 스플링크(Splunk)사가 개발한 Splunk Enterprise이다. Splunk Enterprise는 IT분야에서 발생하는 로그 데이터, 실시간 이벤트 데이터 및 모든 소스의 머신 데이터를 실시간으로 수집·모니터링·검색·분류하고 분석할 수 있는 엔진을 포함하는 솔루션이다. 또한 다양한 분석을 통해 사용자가 원하는 대쉬보드를 자유롭게 생성할 수 있는 운영 인텔리전스(operational intelligence) 플랫폼으로 10여년 전부터 국내외에서 빅데이터 분석도구로 널리 활용되고 있다. Splunk Enterprise는 다음과 같은 기능을 통해 자료에 대한 통찰력과 인텔리전스를 얻을 수 있는 것으로 평가받는다[18].

- 데이터 인덱싱 - 형식이나 위치에 관계없이 모든 머신 데이터를 인덱싱할 수 있다. 사전 정의된 스키마 없이 모든 소스, 형식 및 위치에서 데이터를 인덱싱할 수 있는 기능을 보유하고 있다.
- 검색 및 조사 - 동일한 인터페이스를 사용하여 실시간 및 이력 데이터를 검색할 수 있으며, 통계 보고 명령어를 사용하고, 트랜잭션 개수 업데이트, 매트릭 계산, 일정 시간 윈도우 내에서 특정 조건의 검색도 가능하다.
- 복잡한 이벤트의 상관관계 추적 - 관련 없어 보이는 이벤트나 작업 간의 관계를 쉽게 설정하거나 찾

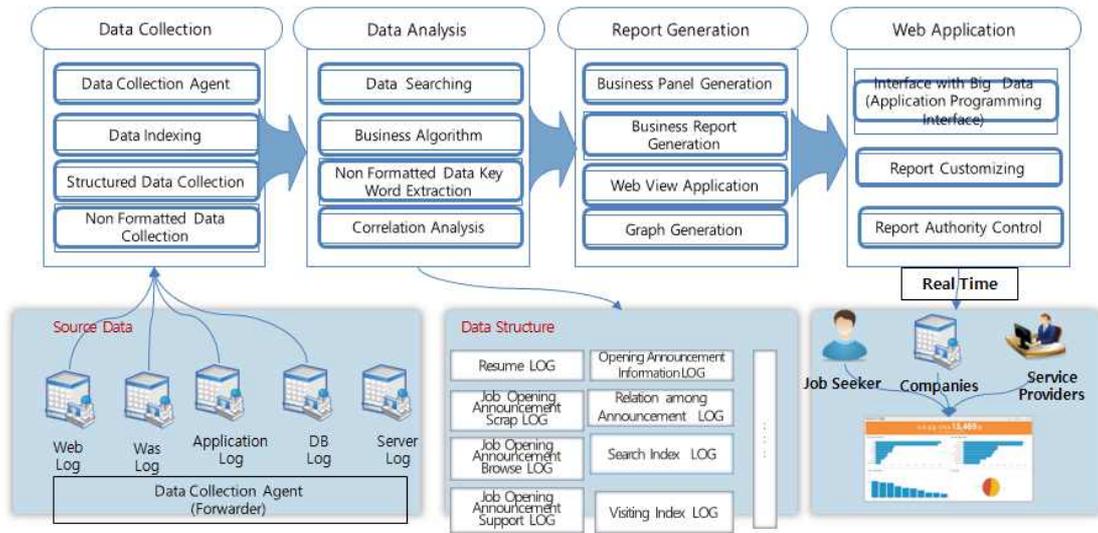


Fig. 4. Data Collection and analysis Process at WorldJob+

을 수 있으며, 시간, 외부 데이터, 위치 등을 기반으로 머신 데이터 내부에 잠재되어 있을 수 있는 상관관계를 도출해낼 수 있다.

- 보고 및 분석 - 보고서, 고급 그래프 및 차트를 작성하여 중요한 동향을 파악하고 고급 시각화를 생성해 주므로 이를 통해 빠른 통찰력을 제공받을 수 있는데, 예측자료의 시각화, 피벗 인터페이스, 대시보드 통합 등의 기능도 제공한다.

4. 데이터 분석사례 결과

본 연구에서는 월드잡플러스를 통해 축적된 구직자 이력서, 공고열람정보, 스크랩한 정보, 공고 지원이력, 사

이트 방문 기록, 검색어 등의 빅데이터 자료를 Splunk Enterprise가 제공하는 다양한 알고리즘 및 상관관계를 통해서 구직자에게 다양한 정보를 제공한다. 빅 데이터 분석을 활성화하기 위해서는 데이터 확보, 데이터 처리 및 저장, 데이터 분석, 정보활용의 4가지 구성요소가 필요하다[19]. 월드잡플러스가 현재 운영 중인 자료분석 프로세스를 간략하게 그림으로 요약하면 Fig. 4와 같다.

월드잡플러스에서 빅데이터 분석을 통해 서비스 이용자에게 제공하는 정보 중 구직자 니즈와 가장 가까운 내용을 제시하는 서비스는 구직자 선호도를 반영한 맞춤형 공고 제공, 구직자와 구입기업간의 맞춤형 매칭 서비스 제공, 구직자를 위한 최적 구직 성공요건 정보제공 등을 들 수 있다. 첫째, 구직자 선호도를 반영한 맞춤형 공고 제공은 다음과 같은 과정을 거쳐 창출되며, 그 과정을 요

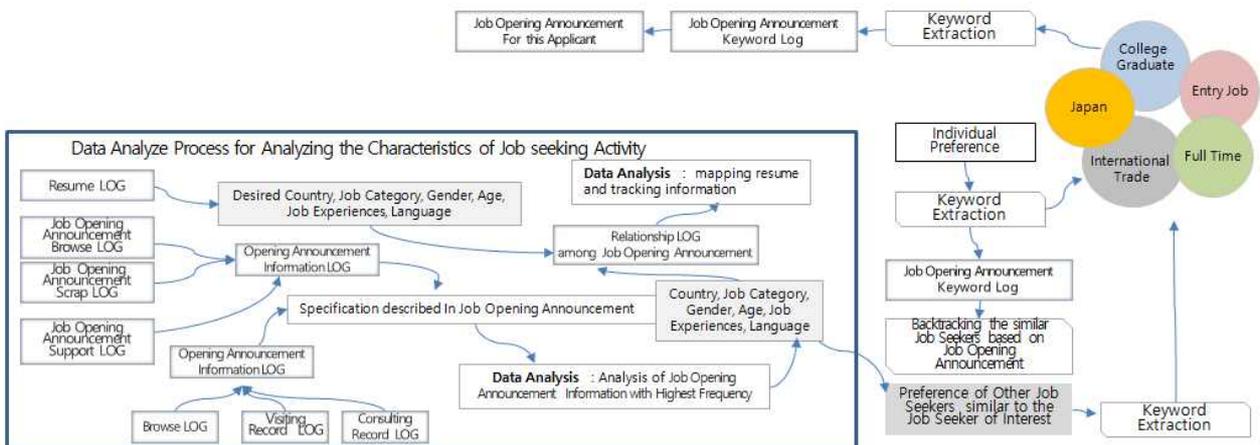


Fig. 5. Matching Job Opening Announcements to the Needs of the Job Seekers

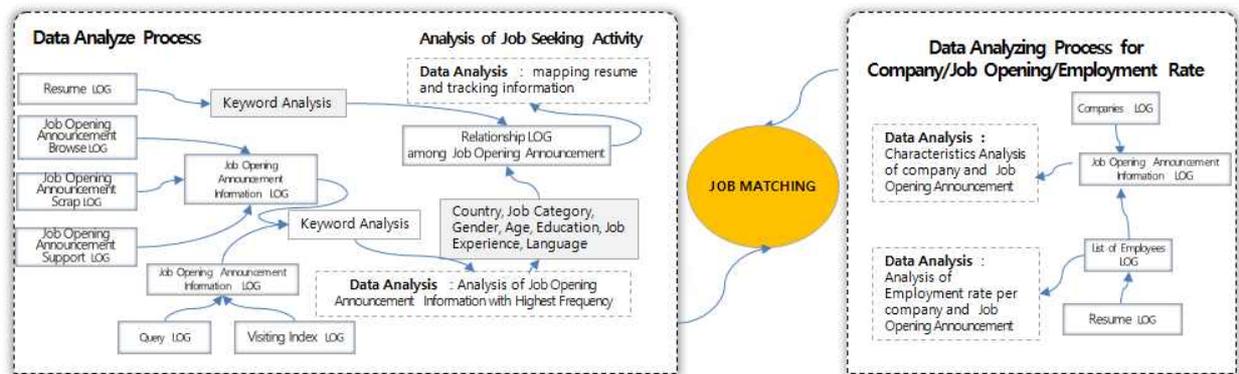


Fig. 6. Matching the needs of Job Seekers to the requirements of the recruiting firm

약하면 Fig. 5와 같다.

- (a) 이력서 상의 희망국가, 직종, 나이, 성별, 학력, 경력, 외국어 구사능력 등에 관한 키워드를 수집한다.
- (b) 또한 이력서와 공고열람, 검색, 방문, 구인공고에 지원한 이력, 스크랩 로그파일에서 국가, 직종 가중치에 따라 구직자 맞춤형 키워드를 추출한다. 더불어 공고지원 로그파일에서 구직자 지원자, 지원일, 합격여부에 관한 정보도 수집한다.
- (c) 키워드 기반으로 구직자가 좋아하는, 구직자 정보와 유사한, 구직자가 좋아할 것 같은, 구직자와 비슷한 이가 좋아하는 구직정보를 분석한다.
- (d) 개인의 선호도와 개인과 유사한 구직자들의 선호도를 토대로 핵심키워드를 추출하여 상관관계 분석을 통한 키워드 기반 맞춤형 해외 진출 공고를 추천한다.

둘째, 구직자와 구인 기업간의 맞춤형 매칭 서비스 제

공은 다음 과정을 거치며, 그 과정을 요약하면 Fig. 6과 같다.

- (a) 맞춤형 공고 제공서비스 분석과정의 첫 세 단계인 구직활동 특성 분석을 통해 희망국가, 직종, 나이, 성별, 학력, 경력, 외국어 구사능력 등의 구직자 키워드를 도출한다.
- (b) 구인공고 정보에서 구인기업의 국가, 직종, 기업형태와 구인공고 정보와 공고에 대한 취업률, 취업성고자의 정보를 분석하여 구직자와 구인기업간 매칭 키워드를 도출한다.
- (c) 사용자의 트래킹을 통해 획득한 구직자 구인활동상태와 구인기업정보 중 매칭 가능한 항목을 도출하여 취업률 등을 증대할 수 있는 매칭결과를 도출하여 구직자와 구인기업 공고간 일대일 해외진출 매칭 정보를 제공한다.

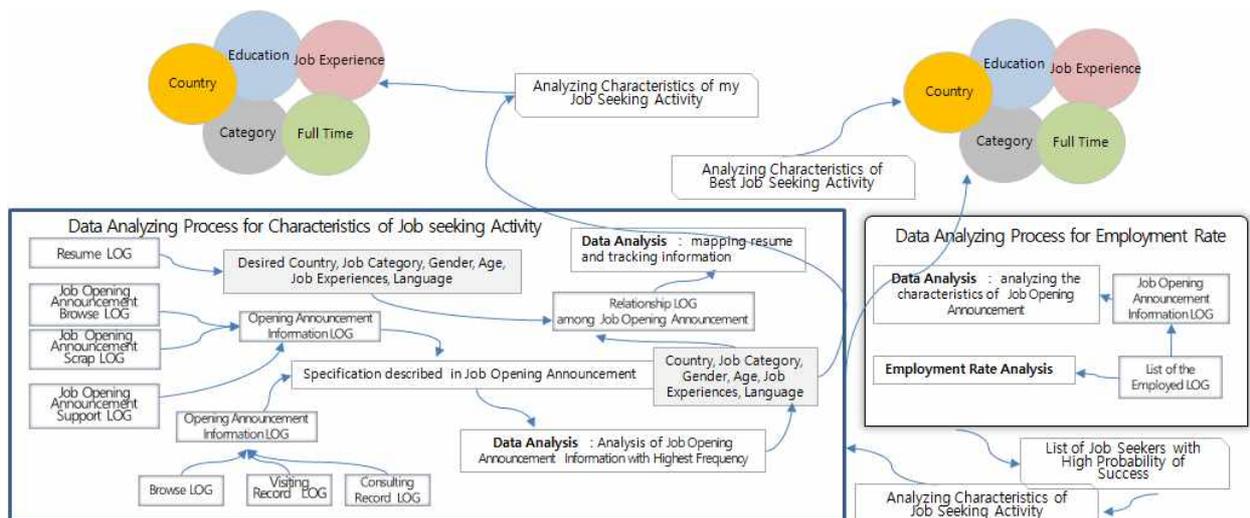


Fig. 7. Finding the Best Specification for Successful Job Seeking Process

마지막으로 성공적인 구직을 위해 구직자에 최적화된 성공요건에 관한 정보는 다음과 같은 과정을 거쳐 도출되며 그 과정을 요약하면 Fig. 7과 같다.

(a) 맞춤형 공고 제공서비스 분석과정의 첫 세 단계인 구직활동 특성분석을 통해 취업에 성공한 구직자들에 관한 희망국가, 직종, 나이, 성별, 학력, 경력, 외국어 구사능력 등의 구직자 키워드를 도출한다.

(b) 취업에 성공한 구직자의 이력서 분석, 공고 열람, 지원 히스토리를 분석하고 나의 구직 이력서, 공고 열람, 스크랩, 로그인 등 구직활동 특성을 분석하여 항목별로 시각화할 수 있는 키워드를 도출한다.

(c) 취업 성공자의 빅데이터와 나의 구직활동 빅데이터를 비교 분석하여 'Best 취업상태'와 '나의 구직 상태'를 인포그래픽스 형태로 제공하여 취업 성공을 위한 방향을 제시하여 청년해외취업을 지원한다.

5. 결론 및 시사점

일반적으로 구직활동은 새로운 노동시장에 진입하기 위해 구직자가 취업하기 원하는 분야의 일자리를 탐색하고, 관련 분야에 대한 상세한 정보를 수집하고, 이를 토대로 자격조건을 갖추어 가는 일련의 행위들로 이해할 수 있다. 월드잡플러스가 제공하는 구직관련정보는 단순히 취업공고 오픈링 목록을 제시하는 정도에 그치는 것이 아니라 빅데이터 분석정보를 통해 적절한 직업의 목록화 및 해당 직업으로의 진입을 위한 실질적인 지원자의 성별, 연령, 학력, 경력 정보 등 예측 정보와 구직자와 구인공고의 키워드 기반 가중치를 분석하여 매칭정보도 제공한다.

이 과정에서 월드잡플러스의 빅데이터 분석정보는 계량화 시킬 수는 없으나 청년구직자들의 구직효능감 제고에 상당한 기여를 할 것으로 기대한다. 구직효능감은 구직활동에서 나타나는 자기관련 유능성으로 구직활동을 성공적으로 수행하여 구직목표를 달성할 수 있다는 믿음을 의미하며[20], 구직자의 구직능력에 대하여 긍정적인 인식을 조성하므로 구직동기를 높여주며 취업 성공에 대한 기대 수준과 몰입수준 제고에도 기여한다. 이들은 구직정보를 탐색하는 범위가 포괄적이며 구직활동에 참여하는 빈도가 높아 공식경로에 대한 정보가 많아지게 된

다. 이와 같은 긍정적 측면을 고려한다면 월드잡플러스를 통한 취업률 개선도 상당할 것으로 전망한다. 월드잡플러스가 빅데이터 분석정보를 제공하기 시작한지 아직 6개월도 되지 않아 구체적인 취업률 개선정도를 추정하기는 어렵지만 기존의 평균 1.3% 정도에 불과했던 지원자들의 취업률이 10% 수준 정도로 제고될 것으로 전망하고 있다.

일자리에 대한 정보가 불확실한 청년들에게 월드잡플러스와 같은 공식경로를 통한 정보제공과 채용지원과정은 상당한 신뢰감을 불러일으키기도 한다. 사회적 네트워크와 같은 비공식적인 경로에 비해 공공기관이 운영하는 공식적 경로는 직장과 직무에 대하여 구체적이고 신뢰성있는 정보를 제공하므로 다양한 취업기회가 제시되어[21], 결과적으로 구직 확률도 높아지게 된다. 청년구직자들이 구직 기간 중 구직활동을 점검하여 자신의 장점을 활용하고 취약점을 보완한다면 자신의 기대에 맞는 직장에 취업할 가능성이 높아진다.

국내 인재들을 원하는 다양한 해외기업들에 관한 구인관련 정보를 수집하고 이들 기업의 비전이나 운영상황에 관한 폭넓은 정보를 입수·분석할 수 있다면 해외취업을 희망하는 국내 구직자들의 니즈를 더욱 효과적으로 충족시켜 줄 수 있을 것으로 기대한다. 월드잡플러스의 빅데이터 분석기능이 예측적, 처방적 역량에는 미치지 못하고 있다는 점을 감안한다면 앞으로 빅데이터 분석을 통해 월드잡플러스가 제공할 수 있는 정보와 서비스의 범주와 품질은 가늠하기 힘든 정도이다. 그렇지만 공공부문 전반에 걸친 빅데이터 분석과 활용은 도입기에 불과하다. 본 연구는 청년들의 해외취업을 지원하는 빅데이터 기반의 정보제공 사례를 주된 내용으로 하고 있지만 공공부문에서 빅데이터 분석이 제공할 수 있는 잠재적 가치를 조명해 볼 수 있었다는 점에서 공학적 연구가 주류인 빅데이터 분야에서 연구의 차별성을 찾아볼 수 있다.

공공부문에서의 빅데이터 분석은 기본적으로 수요자들의 복지적 니즈를 충실히 이해하고 이를 충족시키는데 목적이 있다. 본 연구는 이와 같은 목적을 달성하기 위해 필요한 이론 및 수학적 모형구축과는 다소 거리가 있다. 빅데이터 분석이 가져다 줄 수 있는 막대한 기대감만으로는 공공부문 전반에 걸친 확산이 용이하지 않을 것이다. 본 연구 역시 상용화된 도구를 활용한 분석결과만을 제시한 것이기 때문에 왜 이와 같은 프로세스를 거쳐 데

이 데이터를 분석하고 정보를 가공했는지에 대한 이론적 배경이 충분하지 못하다.

공공부문에서의 빅데이터 분석이 활성화되기 위해서는 공공 데이터 및 정보수요자에 대한 분야별 니즈 분석과 어떤 유형의 빅데이터 분석기반 정보가 요구될 것인지에 대한 정교한 요구분석이 필요할 것이다. 앞으로 이에 대한 후속 연구와 다양한 빅데이터 분석들의 활용방안에 관한 연구가 필요할 것으로 사료된다.

REFERENCES

[1] Gartner. (2017). *Big data*. <http://blogs.gartner.com>.

[2] T. U. Kim. (2016). *Principles of Management*, Seoul : SinYoungSa.

[3] S. H. Lee & D. W. Lee. (2013). Current Status Of Big Data Utilization. *Journal of Digital Convergence*, 11(2), 229-233.

[4] National Informatization Strategy Committee. (2011). *Smart Government with Utilization of Big data*, www.nia.or.kr.

[5] M. Y. Kim & D. J. Seo. (2014). An Analysis of the Public Data for Making the Ambient Intelligent Service. *Journal of Digital Convergence*, 12(12), 313-321.

[6] M. Lee. (2011). Big Data and the Utilization of Public Data. *Internet and Information Security*, 2(2), 47-64.

[7] J. S. Han. (2014). Utilization Outlook of Medical Big Data in the Cloud Environment. *Journal of Digital Convergence*, 12(2), 397-407.

[8] Y. M. Lee. (2017. 4. 8). Big data can locate the welfare dead zone. *E-daily*, p. 12.

[9] NIA. (2015). *Public data becomes the base camp for success*. www.daegu.go.kr/Images/public_data/NIA.

[10] P. Russom. (2011). Big Data Analytics. *TDWI Best Practices Report*, 1-34.

[11] B. Hazen, J. B. Skipper, J. D. Ezell & C. A. Boone. (2016). Big Data and predictive analytics for supply chain sustainability: A theory-driven research agenda. *Computers & Industrial Engineering*, 101, 592-598.

[12] Y. K. Jung, M. Suk & C. Kim. (2014). A Study on the Success Factors of Big Data through analysis of Introduction Effect of Big Data. *Journal of Digital Convergence*, 12(11), 241-248.

[13] J. Huh. (2017). Crime Prevention with Big Data Analysis. *Donga Business Review*, 235, 120-126.

[14] Y. Hahm. (2017). Data Integration Strategy in Big Data

Era: A Public Sector Case Analysis. *Journal of Information Technology and Architecture*, 14(2), 115-128.

[15] S. Kim, H. Shin & S. Son. (2014). A Study on Large-Scale Traffic Information Modeling using R. *Journal of KIISE : System and Theory*, 41(4), 151-157.

[16] B. Y. Lee, J. T. Lim & J. Yoo. (2013). Utilization of Social Media Analysis using Big Data. *Journal of the Korea Contents Association*, 13(2), 211-219.

[17] <https://www.worldjob.or.kr/ovsea/>

[18] <http://www.hancommds.com/splunk/>

[19] B. C. Kim. (2013). Big Data Security Technology and Response Study. *Journal of Digital Convergence*, 11(10), 445-451.

[20] J. Zikic and A.M. Saks. (2009). Job Search and Social Cognitive Theory: The Role of Career-relevant Activities. *Journal of Vocational Behavior*, 74(1), 117-127.

[21] A. Tziner, E. Vered & L. Ophir. (2004). Predictors of job search intensity among college graduates. *Journal of Career Assessment*, 12(3), 332-344.

이 재 덕(Lee, Jae Deug)

[정회원]



- 1993년 4월 : 일본 게이오대학교 이공학연구학부 관리공학과 계산기공학연구실 석사
- 1995년 2월 ~ 1999년 12월 : 삼성항공 정보전략팀 과장
- 2001년 1월 ~ 2016년 12월 : 삼성 SDS 인트라넷 사업부 인트라넷 운영팀장
- 2007년 1월~현재 : 한국산업인력공단 정보화지원국 정보화사업2팀 팀장
- 관심분야 : 빅데이터 분석, 정보공학, 인트라넷운영전략
- E-Mail : jaedeug.lee@gmail.com

이 문 기(Rhee, Moonki)

[정회원]



- 2012년 2월 : 홍익대학교 기계정보공학과(공학학사)
- 2014년 2월 : 성균관대학교 경영학과 MBA(경영학석사)
- 2017년 8월 : 성균관대학교 경영학과(경영학박사)
- 2017년 9월 ~ 현재 : 성균관대학교 경영학과 비전임교수
- 관심분야 : 공급망관리, 빅데이터분석
- E-Mail : mkrhee@skku.edu

김 미 량(Kim, Mi Ryang)

[정회원]



- 1987년 2월 : 서울대학교 인문대학
영어영문학과(문학사)
- 1988년 10월 : 미국 리하이 대학교
교육공학과(석사)
- 1998년 2월 : 서울대학교 교육학과
교육방법/공학 (Ph.D.)

▪ 1999년 3월 ~ 현재 : 성균관대학교 컴퓨터교육과 교수

▪ 관심분야 : SW관리/교육

▪ E-Mail : mrkim@skku.ac.kr