

라이프케어 실현을 위한 빅데이터 활용

신병주, 유성준
세종대학교

요약

적극적인 건강증진과 예방활동을 통해 최적의 건강상태와 높은 수준의 삶의 질을 추구하기 위한 서비스를 제공하는 라이프케어를 실현하려는 보건 의료산업 분야에서도 타 분야와 마찬가지로 증가하는 의료비에 대한 절감 압박, 서비스의 수준에 대한 소비자의 관심 증대 등 당면한 문제 해결과 산업 경쟁력 강화 방안의 일환으로 빅데이터 활용 방안에 대한 논의가 활발히 이루어지고 있다. 이에 본고에서는 라이프케어 실현을 위한 효율적인 빅데이터 활용방안과 이를 위해 해결해야 할 과제와 전망을 제시하고자 한다.

I. 서론

최근 빅데이터(Big Data)는 사회전반에서 새로운 관심영역으로 부상하고 있다. 스마트폰, 스마트TV, RFID, 센서 등의 급속한 보급과 모바일 인터넷과 소셜미디어의 확산으로 데이터 양이 기하급수적으로 증가하고 있고 데이터의 생산, 유통, 소비 체계에 큰 변화를 주면서 데이터가 경제적 자산이 될 수 있는 빅데이터 시대를 맞이하게 됐다. ICT가 다른 산업들과 융복합되면서 방대한 양의 데이터들이 생산되고 있는 가운데 사회변화에 따른 삶의 질에 대한 욕구 및 현안해결에 빅데이터의 활용은 매우 중요한 과제로 떠오르고 있다. 빅데이터는 새로운 가치 창출 엔진으로 인식되면서 산업계 전반과 공공분야, 더 나아가 일상생활에까지 활용분야가 넓어지고 있고, 빅데이터 기반 사회 안전이나 공익 신장의 비즈니스 세계가 새롭게 창조되기도 했다.

한편, 최근 고령화와 국민의료비 지출의 증가와 함께 삶의 질에 대한 관심이 높아지고 있다. 이러한 관심과 사회 트렌드를 일컬어 '라이프케어(Life Care)'라는 개념으로 정의하면서 미래 성장 산업으로 전망하고 있다. 라이프케어란 적극적인 건강증진과 예방활동을 통해 최적의 건강상태와 높은 수준

의 삶의 질을 추구하는데 필요한 서비스를 제공하는 것이다. 보건 의료산업 분야에서는 IT 기술을 융합하여 스마트 헬스케어, U-Healthcare, m-Health 등의 다양한 서비스를 제공하기 위한 노력들이 진행 중이다. 라이프케어는 이와 같은 기존의 보건 의료산업에서 IT 기술을 이용한 서비스들의 범위 확장으로 볼 수 있다. 전문가들은 보건 의료 분야를 진단과 치료 시점의 전과 후로 나누어 구분한다. 치료 시점 전의 건강관리 단계가 웰니스(Wellness)이고 치료가 시작된 이후부터의 관리단계가 헬스케어(Healthcare)이다. 기존의 스마트 헬스케어나 U-Healthcare 등은 치료 중심의 헬스케어 영역을 다루거나 주로 건강관리를 주목적으로 한다. 라이프케어는 웰니스와 헬스케어 영역을 전체를 대상으로 전반적인 삶의 질을 높이는 서비스를 제공하는 것이라고 할 수 있다[10][11][12]. 이처럼 보건 의료산업에서는 라이프케어의 성공적인 실현 뿐만 아니라 증가하는 의료비에 대한 절감 압박, 서비스의 수준에 대한 소비자의 관심 증대 등 당면한 문제 해결과 산업 경쟁력 강화 방안의 일환으로 타 분야와 마찬가지로 빅데이터 활용 방안에 대한 논의가 활발히 이루어지고 있다[1][3][9][13]. 이에 본고에서는 라이프케어 실현을 위한 효율적인 빅데이터 활용방안과 이를 위해 해결해야 할 과제 및 전망을 제시하고자 한다.

II. 본론

1. 빅데이터 개요

빅데이터란 일정시간에 데이터를 처리, 저장, 관리할 때 흔히 쓰이는 소프트웨어의 허용범위를 넘어서는 데이터 덩어리를 가리킨다. 예를 들면 웹 로그, 센서 네트워크, 소셜미디어, 통신 네트워크에서 저장하는 상세 정보나 천체 관측, 생물학적 시스템, 군사 방위, 의학 기록, 사진과 비디오 아카이브 등이다. 빅데이터에 대해 이야기 할 때에는 보통 세가지 측면에서 이야기한다. 첫 번째 양(Volume)의 측면으로 저장할 데이터의 양과

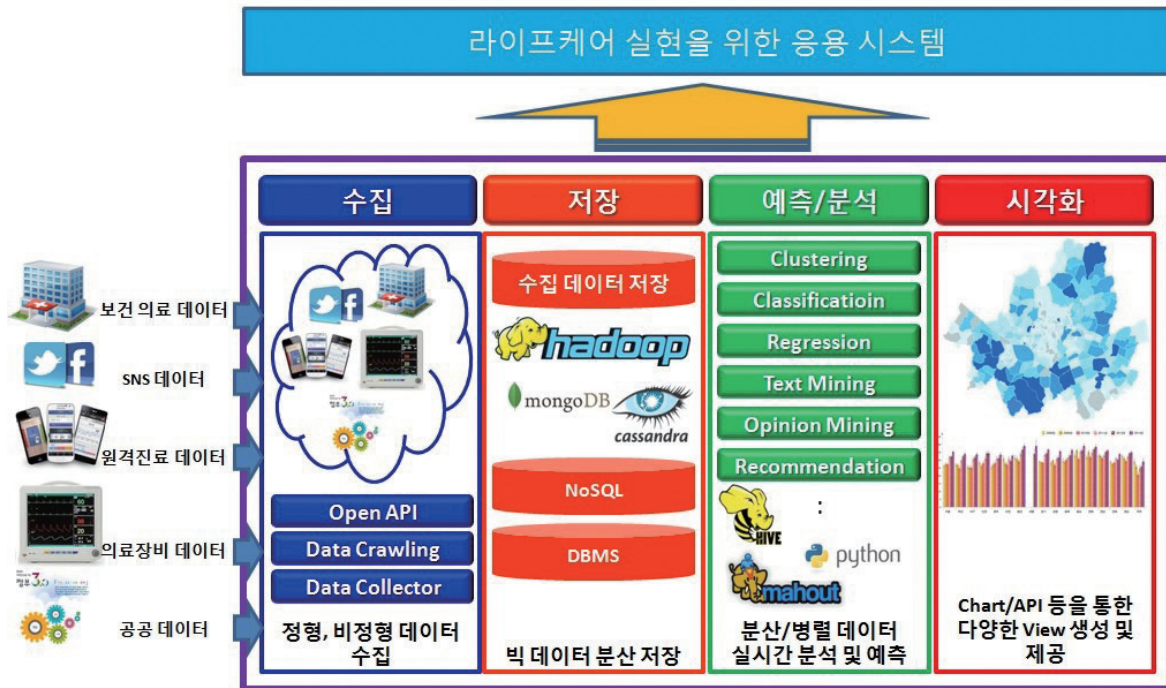


그림 1. 빅데이터 처리 단계

의미 분석과 데이터 가공을 많이 해야 하는 처리 요구량을 말한다. 두 번째 속도(Velocity)의 측면으로 데이터를 처리하는 속도와 저장 속도를 말하며 필요에 따라서 수많은 사용자 요청을 실시간으로 처리한 후 처리 결과를 반환해 주는 기능도 필요함을 의미한다. 세 번째 다양성(Variety)의 측면으로 관계형 데이터베이스에서 사용하는 테이블의 레코드와 같이 정형화되고 사전에 정의할 수 있는 정제된 형태의 데이터 뿐만 아니라 텍스트, 이미지와 같은 비정형 데이터의 처리를 의미한다. 이 세 가지 측면에서 데이터를 분석해야만 가치를 발생시킬 수 있다는 관점에서 가치(Value)의 측면을 추가하기도 한다[2].

2. 빅데이터 처리기술

빅데이터의 분석 기법들은 통계학과, 특히 기계학습, 데이터 마이닝 분야에서 이미 사용되던 기법들이며, 이 분석 기법들의 알고리즘을 대규모 데이터 처리에 맞도록 개선하여 빅데이터 처리에 적용시키고 있다. 최근 소셜미디어 등 비정형 데이터의 증가로 인해 분석 기법들 중에서 텍스트, 오피니언 마이닝, 소셜 네트워크 분석, 군집 분석 등이 주목을 받고 있다.

- **텍스트 마이닝(Text Mining)**: 텍스트 마이닝은 비정형, 반정형 텍스트 데이터에서 자연어처리(Natural Language Processing) 기술에 기반하여 유용한 정보를 추출, 가공하는 것을 목적으로 하는 기술이다. 텍스트 마이닝 기술을 통해 방

대한 텍스트 문치에서 의미 있는 정보를 추출해 내고, 다른 정보와의 연계성을 파악하며, 텍스트가 가진 카테고리를 찾아내는 등, 단순한 정보 검색 그 이상의 결과를 얻어낼 수 있다.

- **오피니언 마이닝(Opinion Mining)**: 오피니언 마이닝은 소셜미디어 등의 정형, 비정형 텍스트의 긍정(Positive), 부정(Negative), 중립(Neutral)의 선호도를 판별하는 기술이다. 오피니언 마이닝은 특정 서비스 및 상품에 대한 시장규모 예측, 소비자의 반응, 입소문 분석(Viral Analysis) 등에 활용되고 있다. 정확한 오피니언 마이닝을 위해서는 전문가에 의한 선호도를 나타내는 표현, 단어 자원의 축적이 필요하다.
- **소셜 네트워크 분석(Social Network Analytics)**: 소셜 네트워크 분석은 수학의 그래프 이론(Graph Theory)에 뿌리를 두고 있다. 소셜 네트워크 연결 구조 및 연결강도 등을 바탕으로 사용자의 명성 및 영향력을 측정하여 소셜 네트워크 상에서 입소문의 중심이나 허브(Hub) 역할을 하는 사용자를 찾는 데 주로 활용된다.
- **군집 분석(Cluster Analysis)**: 군집 분석은 비슷한 특성을 가진 개체를 합쳐가면서 최종적으로 유사 특성의 군(Group)을 발굴하는데 사용된다. 예를 들어 트위터 상에서 주로 사진, 카메라에 대해 이야기하는 사용자군이 있을 수 있고, 자동차에 대해 관심있는 사용자군이 있을 수 있다. 이러한 관심사란 취미에 따른 사용자군을 군집 분석을 통해 분류할 수 있다.



그림 2. 구글 독감예측프로그램

3. 의료분야의 빅데이터 이용 사례

구글은 2008년에 독감예측프로그램[4]을 선보였다. 이 프로그램은 독감과 관련된 검색어의 입력빈도를 분석하여 매우 낮음-낮음-보통-높음-매우 높음의 5단계로 독감의 확산을 예측하는 프로그램이다. 이 프로그램은 2009년 대서양 연안 중부 지역에서 감기가 확산할 것이라고 미국 건강관리본부인 CDC보다 2주 일찍 예측하기도 하였다.

유럽과 미국 등지에서는 이와 같이 의료 분야에서 빅데이터를 이용하려는 사례가 늘고 있다. 미국에서는 의료기관, 환자, 정부, 의료보험회사를 통합해 효율적으로 운영하기 위한 Health 2.0의 일환으로 필박스(Pillbox)[7]라는 의약품 정보 서비스를 제공한다. 미국립 의료원에서 제공하는 이 서비스를 통해 사용 중인 의약품 정보를 제출하는 대신 현재 질병의 전파지역, 전파 속도, 현황과 더불어 연도별 증감을 제공받는다.

IBM에서는 병원업무를 모두 아우를 수 있는 대규모 정보 시스템인 애널리틱스를 위한 IBM 퓨어데이터 시스템(IBM PureData System for Analytic)[6]과 IBM 코그노스10 BI(Cognos 10 BI)[5]를 개발하여 현재 여러 병원에 도입 중이다. 이 프로그램을 통해 병원 운영 및 데이터의 신속한 취합, 정밀한 데이터 분석으로 신속한 의사결정, 연구주제 선정, 논문 작성을 위한 데이터 분석 범위 확장을 할 수 있고 이를 통해 IBM은 가시적으로 의료 애널리틱스를 적극 주도하고 있다.

이외에도 현재 애플이나 삼성과 같은 거대 IT기업들이 관심을 갖고 있는 웨어러블 기기, 애플워치나 삼성 기어와 같은 기기는 헬스 기능을 포함하고 있다. 애플을 예로 들자면, 애플이 자사의 새 운영체제 iOS8에 헬스키트를 기본적으로 탑재한다고 하였다. 이 플랫폼은 각종 헬스케어 기기에서 측정된 데이터를 관리할 뿐만 아니라 실제 의료 서비스까지 연계시키겠다는 목표를 갖고 있다. 이러한 기기가 일상화되고 이러한 플랫폼이 활성화된다면 일상 속에서 IT와 의료가 유기적으로 융합될 수 있는

계기가 될 수 있을 것이다.

4. 국내 의료분야의 빅데이터 도입

분당서울대병원은 의료계 첫 빅데이터 도입이라는 평가를 받고 있다. 2012년 빅데이터 분석 기술을 도입하여 현행 데이터웨어하우스의 문제점을 개선하기 위해 차세대 임상데이터웨어하우스(CDW) 구축 사업을 진행하였다. 이를 통해 진료 기록 처리 시간 단축과 대용량 데이터 분석, 자연어 검색 지원으로 의료진의 연구 검색 활성화, 자료 제공 업무 요청 대기 시간 단축을 통한 효율성 증대 등의 효과를 거둘 수 있을 것으로 보고 있다.

전자무기록시스템(ERM) 도입 확산으로 주요 대형병원들을 주축으로 데이터웨어하우스(DW)에 대부분 형식에 맞춰 입력된 정형화된 데이터 중심으로, 대량의 정보를 분석하고 활용하는데 한계가 따른다. 또한 의료진들이 전자차트에 작성한 내용 가운데 비정형 데이터의 검색과 통계가 힘들어 실제 활용에는 제한되어 있다[8].

Ⅲ. 결론

보건의료분야는 예방 건강관리 중심으로 패러다임 전환에 따른 적응과 더불어 국민의료비 증가, 병원 수익구조 악화 등 어려움을 극복하기 위한 산업계의 전반적인 효율성 제고 노력이 이뤄져야 하는 상황이다. 이같은 문제점을 극복하고 산업 경쟁력을 확보하기 위해 주목받고 있는 빅데이터의 활용은 필수가 되고 있다. 이미 보건의료분야에 빅데이터를 활용한 선진국들은 공공부문의 데이터 공개를 통한 사회후생 증가, 기반 인프라 확충을 바탕으로 한 민간참여 유도 정책을 시행하고 있다. 우리나라의 경우 세계적인 수준의 ICT 기술과 인프라를 보유하고, 빅데이터 활용 잠재력 측면에서 경쟁우위에 있다. 하지만 아직까지 보건의료분야에서 빅데이터 활용 초기단계이기 때문에 적극적으로 빅데이터 활용을 통한 부가 가치 창출을 이끌어내기 위해 벤처 및 중소기업 등 빅데이터 활용에 대한 아이디어는 있으나 인프라가 취약한 기업에 데이터 분석 서비스를 저렴하게 제공하거나 빅데이터 활용 노하우에 대한 교육 및 컨설팅 서비스를 제공할 수 있도록 해야 한다. 또 빅데이터 활용 사례를 축적해 지속적인 노하우를 제공하면 기업 및 우수 인재들의 보건산업 빅데이터 분야 유입을 촉진할 수 있을 것으로 보인다. 빅데이터 활용에 있어 미비한 법과 제도의 정비, 개인정보에 대한 보안 기술 및 인식 개선, 데이터의 집중에 따른 권력화 등 해결해야 할 과제들은 아직 많이 남아있다. 새롭게 생겨나는 분야인 만큼 충분한 해외 사례 벤치마킹을 통해 예상되는 부작용을

최소화하면서 민간부문의 혁신역량을 활용할 수 있는 법과제도 체계의 정보가 요구된다. 또 무분별한 데이터 활용을 방지하기 위해 데이터 수집과 활용의 용도와 범위, 절차 및 타당성에 대한 가이드라인 제정이 선행돼야 할 것이고, 데이터 독점에 따른 감시자 출현, 권력화 현상 발생 등 예상되는 사회적 부작용에 대한 검토와 논의가 충분히 이뤄져야 할 것으로 보인다.

참고 문헌

- [1] Chawla, N.V. and Davis, D.A. "Bring Big Data to Personalized Healthcare: A Patient-Centered Framework", *Journal of General Internal Medicine*, Volume 28, Issue 3, pp. 660-665, 2013
- [2] Chen, M, Mao, S., and Liu, Y. "Big Data: A Survey", *Journal of Mobile Networks and Applications*, Volume 19, Issue 2, pp. 171-209, 2014
- [3] Sun, J. and Reddy, C.K. "Big Data Analytics for Healthcare", *Proceedings of the 19th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 1525, 2013
- [4] Google.org 독감 트렌드. (<http://www.google.org/flutrends/intl/ko/>)
- [5] IBM Cognos 10.2 BI Demo. (http://www.ibm.com/New_Cognos10.2_Demo)
- [6] IBM PureData System for Analytics. (<http://www.ibm.com/software/data/puredata/analytics>)
- [7] Pillbox - National Library of Medicine. (<http://pillbox.nlm.nih.gov>)
- [8] 김성수. "의료기관의 빅데이터 활용방안에 대한 연구", *디지털융복합연구* 제12권 제2호, pp. 397-407, 2014
- [9] 이지혜, 제미경, 조명지, 손현석. "보건의료 분야의 빅데이터 활용 동향", *한국통신학회논문지* 제32권 제1호, pp. 63-75, 2014
- [10] 정운수, 퍼지 AHP를 이용한 헬스케어 환자의 빅 데이터 사용의 효율적 관리 기법,
- [11] 조영복, 우성희, 이상호. "웰니스를 위한 빅데이터 분석과 의료 질 관리", *한국컴퓨터정보학회논문지* 제19권 제12호, pp. 101-109, 2014
- [12] 조위덕, 양승국, 최선탍, 백재순, 민명기, 이영권, 박경찬, 이규필. "IoT를 사용한 라이프로그 빅데이터기반 라이프스타일 (생활패턴) 분석 및 웰니스 예측케어 서비스시스템",

- 한국통신학회논문지 제31권 제12호, pp. 17-24, 2014
- [13] 한정수. "클라우드 환경에서 의료 빅데이터 활용 및 전망", *디지털융복합연구* 제12권 제6호, pp. 341-347, 2014

약 력



신 병 주

2001년 경남대학교 공학사
 2003년 경남대학교 공학석사
 2008년 경남대학교 공학박사
 2009년~2013년 Research Professor, Brigham Young University, U.S.A.
 2014년~현재 세종대학교 컴퓨터공학과 연구교수
 관심분야: 빅데이터, 텍스트마이닝, 데이터베이스, 정보검색



유 성 준

1982년 고려대학교 공학사
 1990년 고려대학교 공학석사
 1996년 미국 시라큐스대학교 공학박사
 2014년~현재 빅데이터산업진흥센터장
 2002년~현재 세종대학교 컴퓨터공학과 교수
 관심분야: 빅데이터, 데이터마이닝, 텍스트마이닝, 데이터베이스