

의료 IT와 빅데이터의 융합

- 김관형 (동명대학교 컴퓨터공학과)
- 이준연 (동명대학교 미디어공학과)
- 오암석 (동명대학교 미디어공학과)

I. 서론

오늘날 국내 총 인구에서 65세 이상 고령자가 차지하는 비율은 12.2%로 1970년 3.1%에서 지속적으로 증가하여 2030년 24.3%, 2050년 37.4% 수준에 이를 것으로 전망되고 있고 있으며 2012년 건강보험의 고령자 진료비는 16조 382억 원으로 1인당 고령자 진료비가 전년에 비해 2.5% 증가한 293만원으로 매년 증가하는 추세이다.[2,3]

표 1. 고령자 진료비 현황

년도	65세 이상 진료비 (백억원)	1인당 고령자 진료비 (천원)
2005	606	1,545
2006	739	1,815
2007	908	2,070
2008	1,049	2,281
2009	1,204	2,495
2010	1,378	2,769
2011	1,484	2,862
2012	1,604	2,933

이와 같이 사회 고령화에 따른 의료비 지출에 대한 부담이 증가하고 의료서비스의 접근성 및 질에 관한 문제가 제기되면서 의료분야에서는 IT기술을 접목하여 의료서비스의 품질을 높이고, 의료비용을 절감시켜려는 노력이 지속적으로 이

루어져 왔다.[4] 그중 유헬스는 의료비 절감 등의 사회경제적 비용감소 효과와 공공보건 의료서비스와 예방관리 보건 등의 사회정책적 효과를 기대할 수 있는 효과적인 대안으로 각광 받아 왔으며 최근에는 스마트폰, 태블릿 PC 등 다양한 스마트 기기의 보급 확산과 유무선 통신기술과 센싱 기술의 발전으로 스마트 시대가 도래하면서 보다 복합적이고 지능화된 의료서비스가 가능하게 되었다. 이러한 의료서비스의 발전으로 의료분야에서의 데이터 활용은 환자 진료기록 데이터뿐만 아니라 임상 데이터, 의료 영상 이미지, 유전자 통계, 질병 정보, 생활정보, 환경정보 등으로 범위가 점점 확대되고 있으며 이를 활용한 다양한 의료서비스가 등장하고 있다.[5]

하지만 현재 의료서비스 및 기술의 발전으로 서비스 제공자가 데이터를 모두 확인할 수 없을 정도의 빠른 속도로 각종 센서와 모니터, 의료기기 등 다양한 부분에서 방대한 양의 의료 데이터가 발생하고 있어 기존 데이터 저장·관리·분석기법의 한계에 직면해 있다. 이러한 문제로 인해 최근 의료산업에서 실시간으로 발생하는 방대한 양의 데이터들의 신속한 처리와 관리를 위한 해결 방안으로 빅데이터의 활용이 매우 중요한 과제로 부각되면서 다양한 의료분야에서의 빅데이터 관리 및 활용에 대한 연구가 활발히 진행되고 있으며 이를 통한 의료IT와 빅데이터의 융·복합을 통한 새로운 형태의 의료서비스 및 의료비 절감 효과를 기대하고 있다.[1,6]

먼저 빅데이터의 개념과 현황을 살펴보고 의료IT와 빅데이터의 융·복합 사례를 분석하여 의료분야에서 빅데이터의 효율성에 대해 고찰하였다.

II. 관련 연구

1. 빅데이터의 개념

빅데이터는 일반적으로 데이터베이스 시스템으로 관리하기 어려운 정도의 수십 테라에서 페타바이트(Petabyte), 엑사바이트(Exabyte) 정도 크기의 방대한 규모의 데이터를 의미하며 최근에는 대용량 수집, 저장, 분석, 체계화를 위한 도구 및 플랫폼, 분석기법 등을 포괄하는 용어로 변화하고 있다. 이처럼 빅데이터가 포괄적인 용어로 사용되면서 빅데이터에 대한 정의와 특징은 다양하게 규정되어 왔는데, IBM에서는 오늘날 인류는 매일 2.5 쿼틸리언 바이트의 데이터를 생산하는데, 기상정보를 수집하는 센서, 소셜미디어 사이트의 웹문서, 디지털사진과 동영상, 구매거래기록, 휴대전화의 GPS 신호 등 모든 곳으로부터 생성되는 데이터를 빅데이터라 정의하고 있으며 빅데이터는 데이터 크기(volume), 데이터 전달 속도(velocity), 데이터의 다양성(variety), 정확성(veracity) 등 4V로 이루어진 4차원적 특징을 가지며 단순히 크기와 관련된 문제가 아니고, 새로운 데이터처리 및 분석 방법을 통해 새로운 통찰력을 찾을 수 있는 기회를 의미한다고 정의하고 있다. 또한 Gartner에서는 2001년, 데이터 크기가 급증하고(volume), 데이터 전달 속도가 빠르며(velocity), 데이터 구조가 다양한(variety) 현상을 관찰하여 이를 3V로 표현하였고, 2012년 빅데이터를 “크기가 크고, 속도가 빠르며, 다양한 정보자산을 가지고 있다. 이는 새로운 데이터처리방법을 필요로 하는데, 이를 통해 새로운 통찰력의 발견이 가능해진다”라고 표현하고 있으며 MIKE2.0에서는 빅데이터의 가장 중요한 요소는 데이터의 크기이지만, 보다 정확하게는 독립적 데이터 소스 사이의 상호작용 또는 연관관계의 크기를 의미하며 데이터 소스 사이의 연관관계가 복잡하여 데이터 정제와 유의미한 데이터만 추출해내는 것이 어렵다는 점에서 빅데이터의 ‘빅(big)’은 단순히 크기(big volume)가 아니라 복잡성(big complexity)에 대한 것으로 해석하는 것이 적절하다고 말하고 있다.[7,8,9]

즉, 위 내용을 종합해보면 빅데이터는 데이터의 엄청난 크기, 빠른 전송, 다양성 및 비정형성, 복잡성과 같은 특징을 가지는 다루기 어렵고, 새로운 지적 통찰력을 제공할 수 있는 데이터로 정의할 수 있다.

2. 빅데이터의 현황

오늘날 스마트폰의 보급률이 3474만 명으로 전체 이동전화 가입자 중 64.5%를 차지하는 등 스마트폰의 급속한 보급과 다양한 스마트기기의 등장으로 모바일, 온라인 상거래, 소셜네트워크 서비스를 통해 발생하는 데이터량이 기하급수적으로 증가하고 있으며 이렇게 생산된 방대한 양의 데이터에 대한 저장장치의 발달, 초연결성의 확대, 처리기술의 발로 인해 빅데이터 시대로 진입하고 있다.[10] 이러한 빅데이터의 부각으로 방대한 데이터를 통한 가치창출이 기업뿐 아니라 국가의 경쟁력 강화와 직결되면서 빅데이터가 차세대 산업을 이끌어 갈 핵심동력으로 주목받고 있으며 기업과 정부에 새로운 발전의 기회를 제공하거나 기존의 다양한 산업에서의 문제 및 한계를 극복할 수 있는 단초를 제공할 것으로 기대하고 있다.

표 2. 빅데이터 세계시장 규모

년도	규모(백만 달러)
2010	3,217
2011	4,766
2012	6,842
2013	9,728
2014	12,941
2015	16,920

또한 시장조사 전문기관인 IDC에서는 ‘전 세계 빅데이터 기술 및 서비스 전망보고서’를 통해 표 2와 같이 전 세계 빅데이터 시장이 2010년 32억 달러에서 오는 2015에는 169억 달러 규모에 달할 것이고, 전체 정보통신기술(ICT) 시장 성장률의 약 7배에 달하는 연평균 40% 성장률을 보일 것으로 전망하고 있으며 Mckinsey에서는 빅데이터 활용 시 미국의료분야에서 연 3,000억 달러, 유럽 공공분야에서 연 2,500억 달러의 경제적 효과가 발생하고 우리나라의 경우 약 10.7조의 정부지출을 감소시킬 수 있으며 일본의 경우 부가가치의 창출이나 사회적 비용의 절감에서 총 16조원 이상의 경제적 효과가 얻어질 것으로 예상하고 있다.[11,13]

표 3. 주요국 빅데이터 관련 정책 현황

국가	추진기관	세부내용
미국	과학기술 정책실	<ul style="list-style-type: none"> · '빅데이터 연구개발 이니셔티브' 발표 (2012, 3) · 부처별, 지방정부별 빅데이터 활용한 서비스 발굴 및 운영 · 공공정보 데이터 공개사이트 'data.gov' 운영
영국	기업혁신 기술부	<ul style="list-style-type: none"> · '데이터 전략위원회(Data Strategy Board)' 설립(2012, 3) · 데이터 공유플랫폼 'data.gov.uk' 운영
싱가포르	경제개발청	<ul style="list-style-type: none"> · 국가안보조성국(NSCS) 내 RASH 시스템 마련(2004, 7) · 민간협력으로 데이터분석연구소 설립 · 공공정보 데이터 공개사이트 'data.gov.sg' 운영
대한민국	국가정보화 전략위	<ul style="list-style-type: none"> · '빅데이터를 활용한 스마트정부 구현방안' 마련(2011, 10) · 빅데이터 마스터플랜 추진 및 빅데이터 국가전략포럼 실시 · 공유자원포털 'data.gov.kr' 운영

이에 표 3과 같이 세계 각국에서는 의미있는 정보를 실시간으로 도출하거나 트렌드 파악, 마케팅, 의사결정 등 다양한 분야에서 빅데이터를 활용하기 위하여 각 분야에 걸쳐 데이터 분석에 의한 예측과 서비스 개발을 위한 정책 추진 및 지원을 적극적으로 진행하고 있으며 글로벌 리서치 기관, 컨설팅 그룹 등 글로벌 기업에서는 차세대 키워드로 빅데이터를 선정하고 경제적 가치에 주목하여 정보수집과 분석을 위한 대용량 데이터 처리 기술 및 솔루션을 중심으로 조직을 재편하고 역량 강화 및 기술 개발을 위해 노력하고 있다.[12,14]

표 4. 글로벌 기업들의 빅데이터 관련 현황

기업명	내용
EMC	<ul style="list-style-type: none"> · 데이터 저장부터 관리, 분석까지 빅데이터와 관한 모든 것을 제공하기 위해 그린플름, 아이실론 등 빅데이터 솔루션 업체 및 데이터 관련 다수 업체를 인수 · 빅데이터 스토리지 솔루션, 콘텐츠 관리 솔루션 등 제공 'EMC 애널리틱스 랩'을 운영하여 데이터 사이언티스트 육성
IBM	<ul style="list-style-type: none"> · 지난 5년간 140억 달러 이상을 투자하여 분석용 데이터 저장 관리 업체, 데이터 통합 업체, 분석 솔루션 업체 등 비즈니스 분석 관련 다수업체를 인수 · 지속가능한 지구를 만들기 위해 지구 데이터(기온, 토양상

	태, 교통 흐름 등)를 분석하는 '스마트 플래닛' 프로젝트 전개
오라클	<ul style="list-style-type: none"> · 세계적인 DB업체로 '하이퍼리온'시를 인수해서 분석 기술을 확보 · 빅데이터 어플라이언스 제품 출시
SAP	<ul style="list-style-type: none"> · 업무용 어플리케이션 업체에서 최근 DB 전문 업체로 변화 · 메모리 기반 DB 어플라이언스 제시 · IT 소프트웨어, 플랫폼을 제공하는 '비즈니스 오브젝트'사를 인수
테라 데이터	<ul style="list-style-type: none"> · 데이터웨어하우징 및 비즈니스 인텔리전스(BI) 전문 업체 · 비정형 데이터의 고급분석·관리 솔루션 업체 인수 · '에스터 맵리뷰스 플랫폼' 제시
HP	<ul style="list-style-type: none"> · BI 솔루션 업체 '버티카', 기업용 검색 엔진 업체 '오토노미'를 인수 · 버티카와 오토노미를 결합하여 빅데이터 분석 시장에 진입 · '인스턴스온 엔터프라이즈' 솔루션으로 기업경영 의사결정, 경영정보 분석 등 경영지원 전략 수립 서비스 제공
마이크로소프트 레티지	<ul style="list-style-type: none"> · 비즈니스 인텔리전스(BI) 소프트웨어 공급업체 · BI에 빅데이터 분석 처리를 접목하여 사업 역량을 강화
구글	<ul style="list-style-type: none"> · GFS(Google File System), MapReduce, Bigtable 등 대용량 데이터 처리 기술을 발표 · 2011년 이용자 업로드한 대용량 데이터 분석 처리를 지원하는 빅쿼리 서비스 공개
MS	<ul style="list-style-type: none"> · 윈도우 애저와 윈도우 서버 플랫폼용 아파치 허둠 개발 계획
넥스알	<ul style="list-style-type: none"> · 빅데이터 관리 솔루션인 '넥스알 빅데이터 어널리틱스 플랫폼' 출시

III. 의료분야의 빅데이터 융합 사례 및 특징

기존 의료분야에서 빅데이터의 활용은 주로 유전자 염기서열 분석 등의 의학 연구 분야에 집중되어 있었다. 하지만 최근 빅데이터 분석 방법이 발전하고 비용이 낮아지면서 유전공학 외에도 고혈압, 당뇨 등 보다 실제적인 치료를 위한 의료시스템에 활용되기 시작했으며 최근 들어 환자의 의무기록, 식생활 습관, 직업 이력 등 보다 광범위한 데이터의 분석을 통해 치료 방법을 개선하고 의료비 부담을 줄이려는 노력이 진행되고 있다.

이에 본 장에서는 의료분야에서 실제 빅데이터의 활용 사례들을 분석하고 이를 통해 빅데이터 융합에 대한 효율성을 살펴보았다.

1. 필박스(Pillbox)



그림 1. 필박스(<http://pillbox.nlm.nih.gov/>)

필박스는 미국 오바마 정부의 국립보건원에서 환자, 의료 기관 및 보험회사 간의 유기적 연결을 통한 정책 마련 및 비용 절감을 위해 민간 웹 서비스에서 다양한 아이디어를 차용하고 매우 직관적인 인터페이스를 활용하여 누구나 쉽게 접근하여 사용자의 의약품 오남용을 막기 위해 마련된 헬스케어 시스템으로 일약 제조사와 사용자 간의 유기적인 쌍방향 상호작용을 통해 약에 대한 정보를 제공하여 사용자가 사용 중인 약에 대한 정보가 불분명할 때 필박스를 통해 약에 대한 정확한 정보의 확인이 가능하고 알약에 새겨진 글자, 번호, 색깔, 모양, 크기 등 간단한 약에 대한 설명만으로도 정확한 약의 효능 및 정보를 제공 받을 수 있다.

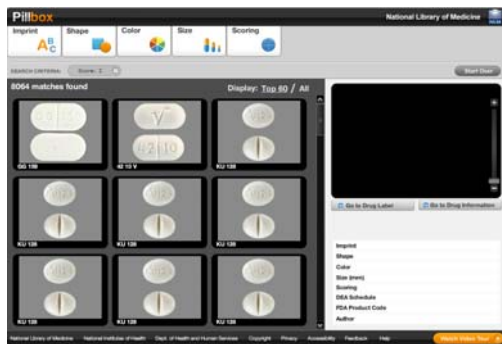


그림 2. 필박스 의약품 조회 기능 1

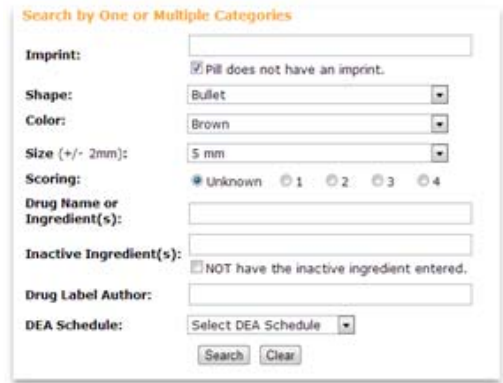


그림 3. 필박스 의약품 조회 기능 2

이에 국립보건원은 한해 동안 접수되는 알약의 기능이나 유효기간을 문의하는 민원 수의 100만 건 이상 필박스로 제공하고 있으며 평균 한 건당 알약의 기능 및 유효기간을 확인 하는데 필요한 비용인 50달러를 필박스 서비스를 제공함으로써 연간 5,000만 달러의 비용을 절감하고 있다.[10,16]

또한 국립보건원에서는 빅데이터를 이용한 의약품 사용에 대한 정보를 제공함으로써 사용자가 검색한 약의 정보를 통해 의약품 사용과 관련된 지도 및 그래프를 입체적으로 작성하여 현재 유행하고 있는 질병의 발생 장소 및 전염 속도에 대한 분석이 가능해졌고 사용자에 의해 만들어진 데이터를 통해 질병에 대한 인적대응 방법, 방제인원에 대한 효율적인 방제대책 마련이 가능하게 되었다.

2. 구글 플루 트렌드

구글에서는 무선랜(WiFi) 이용지역이 확대되고 스마트폰, 태블릿 PC 등 다양한 스마트 디바이스의 빠른 확산으로 편리한 인터넷 사용 환경이 마련되면서 사용자들의 현재 상태나 이에 따른 후속 상황을 예측할 수 있는 정보 등 실시간으로 무수히 많이 누적되는 정보 속에서 사회적인 변화나 흐름을 파악하는 것이 가능해지면서 구글 홈페이지에서 독감, 인플루엔자 등 독감과 관련된 검색어 쿼리의 빈도를 조사하여 독감 확산에 대한 정보를 제공하기 위해 ‘구글 플루 트렌드’라는 경보체계를 마련하였다.[18,19]

주 | 도시(예상치)

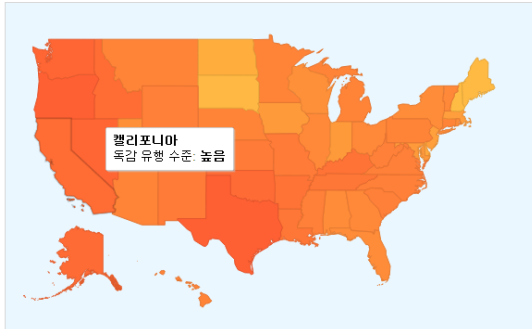


그림 4. 구글 플루 트렌드

또한 구글은 미국 질병통계 예방 센터의 데이터와 비교한 결과, 검색 빈도와 실제 독감증세를 보인 환자의 숫자 사이에 매우 밀접한 상관관계가 있는 것을 확인함으로써 미국 보건 당국보다 한발 앞서 시간 및 지역별 독감 유행 정보를 제공할 수 있게 되었으며 이를 토대로 세계 곳곳에서 다양한 사용자에게 검색어의 분석을 통해 정치, 문화, 경제, 스포츠등 사용자의 관심분야에 대한 유의미한 데이터를 가공하여 제공하고 있다.



그림 5. 구글 플루 트렌드와 실제 데이터 비교 결과

3. 웰포인트의 헬스케어 서비스

고령인구에 의한 의료비가 국내의 경우 2009년 이미 전체 의료비의 30%를 넘어섰고 선진국의 경우 전체 의료비의 40~50%를 차지하는 등 당뇨병, 고혈압, 심장질환과 같은 만성 질환에 의한 의료비 부담이 GDP의 6%를 넘어가면서 병원 정보화를 통해 의료서비스의 효율을 극대화하는 디지털병원에 대한 관심이 증가하고 있다.

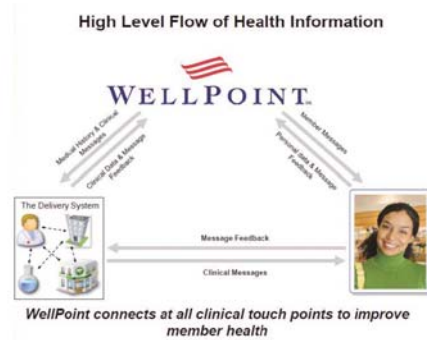


그림 6. 웰포인트의 의료정보 제공 체계

이에 미국 의료보험사인 웰포인트(WellPoint)는 IBM의 클라우드 컴퓨팅 방식의 인공지능 컴퓨터 왓슨을 활용해 건강보험 자료와 회사에 등록된 3,420만 명에 대한 환자의 증상, 환자 면담 결과, 진단 연구 등 진료 내역에 대한 모든 정보를 저장하고 있으며 IBM 서버를 통해 임상실험 결과 및 베스트 프랙티스 같은 과거 사례를 분석하여 최적의 진단 및 환자 치료 가이드라인을 제시하고 있다.[10,17]

이를 통해 웰포인트는 환자의 상황에 맞는 가장 최선의 치료 방법을 제시함으로써 불필요한 치료 및 진료를 줄여 환자 및 의료보험 회사의 불필요한 진료비 낭비를 방지하고 있으며 모든 정보 및 사례의 분석을 통한 효과적인 진료서비스의 제공으로 헬스케어의 진단을 및 치료율 향상에 기여하고 있다.

4. 1000 유전자 프로젝트

미국에서 한 해 220만 명 이상이 약물 부작용으로 입원하는 등 인간의 35,000개의 유전자에 따라 같은 질환이라도 증상이 다르게 나타나며, 약물의 반응도 다르게 발생한다. 이에 미국 국립보건원은 75개 기업 및 기관들과의 파트너십을 통해 공동으로 진행한 1000 유전자 프로젝트(1000 Genomes Project)의 일환으로 200TB의 유전자 정보를 확보하고 인종별, 국가별 특성에 따른 유전자 특성 정보를 분석하여 질병 치료에 활용하고 있다.

또한 국립보건원은 1000 유전자 프로젝트를 아마존 웹 서비스(AWS)로 이전하여 저장하고 있으며 아마존 웹 서비스에서는 익명의 개인들로부터 유전자 정보를 제공받아 현재 1,700명의 유전자 정보를 아마존 클라우드에 저장하고 있으며 이러한 유전자 정보의 공유 및 분석은 새로운 질병에 대한

빠른 진단 서비스의 제공을 가능하게 하고 난치병 및 불치병과 관련된 새로운 치료제 개발 가능성을 제시함으로써 질병에 대한 치료 확률을 높일 수 있는 기회를 마련해준다.[10,20]

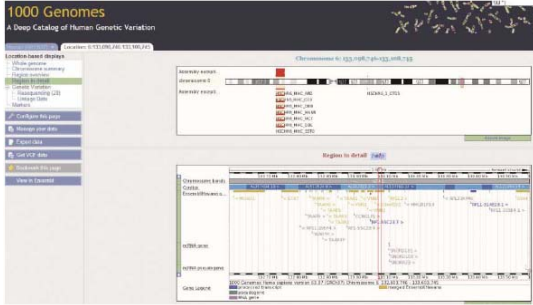


그림 7. 1000 유전자 프로젝트 데이터 검색

1000 Genomes Samples						
Population	DNA sequenced from blood	Offspring Samples from Trios Available	Pilot Samples	Phase 1 Samples	Final Phase Sample	Total
Chinese Dai in Xishuangbanna, China(CDX)	no	yes	0	0	99	99
Han Chinese in Beijing, China (CHB)	no	no	91	97	103	106
Japanese in Tokyo, Japan (JPT)	no	no	94	89	104	105
Kinh in Ho Chi Minh City, Vietnam (KHV)	yes	yes	0	0	101	101
Southern Han Chinese, China (CHS)	no	yes	0	100	108	112
Total East Asian Ancestry (ASN)			185	266	515	523
Bengali in Bangladesh (BEB)	no	yes	0	0	86	86
Gujarati Indian in Houston, TX (GIH)	no	yes	0	0	106	106
Indian Telugu in the UK (ITU)	yes	yes	0	0	103	103
Punjabi in Lahore, Pakistan (PJL)	yes	yes	0	0	96	96
Sri Lankan Tamil in the UK (STU)	yes	yes	0	0	103	103
Total South Asian Ancestry (SAN)			0	0	494	494
African Ancestry in Southwest US (ASW)	no	yes	0	61	66	66
African Caribbean in Barbados (ACB)	yes	yes	0	0	96	96
Esan in Nigeria (ESN)	no	yes	0	0	99	99
Gambian in Western Division, The Gambia (GWD)	no	yes	0	0	113	113
Luhya in Webuye, Kenya (LWK)	no	yes	102	97	101	116
Mende in Sierra Leone (MSL)	no	yes	0	85	85	85
Yoruba in Ibadan, Nigeria (YRI)	no	yes	106	88	109	116
Total African Ancestry (AFR)			208	246	669	691
British in England and Scotland (GBR)	no	yes	0	89	92	94
Finnish in Finland (FIN)	no	no	0	93	99	100
Iberian populations in Spain	no	yes	0	14	107	107

그림 8. 1000 유전자 프로젝트 샘플

5. Virtual Autopsy Table

최근까지 CT 촬영을 통해 얻을 수 있는 3만 개의 신체 영상 모두를 관리 또는 활용할 수 없기 때문에 일부 단면만을 촬영하고 있으나 빅데이터 기술을 활용하면서 환자의 몸 전

체를 촬영한 데이터로 가상의 인체를 구성하는 것이 가능해지고 있다.

이러한 기술은 가상 부검(Virtual Autopsy)라 불리우며, 골격뿐 아니라 장기, 근육, 피부 등 모든 부분을 겹겹이 촬영하여 하나로 연결하고 분석함으로써 어느 부위가 어떻게 손상되었는지 모든 방향에서 높은 해상도로 정확하게 파악할 수 있다.[9,11]

Virtual Autopsy Table은 이러한 가상 부검을 위해 환자의 몸 전체를 촬영한 데이터로 가상의 인체를 구성하는 기술로 환자의 질병 또는 손상 부위를 파악하여 가상의 환경에서 임의의 치료를 통해 정확한 치료법을 찾을 수 있으며 해당 치료법에 대한 부작용의 위험을 낮출 수 있다.



그림 9. Virtual Autopsy Table

6. 무선 혈당 측정 시스템

빅데이터 분석 방법의 발전으로 다양한 의료 시스템에 빅데이터가 활용되면서 최근에는 스마트폰을 이용해 지속적으로 심박수, 혈당, 혈압과 같은 신체 데이터를 측정하고 결과를 전송하여 실시간으로 자신의 상태를 피드백 받는 것이 가능하다.[10]



그림 10. TelCare의 무선 혈당측정 시스템

통계기준: 2012.12.31.

코호트별 인체자원 누적 수집현황

코호트	누적인원*	자원종류별 수집 바이알			
		Serum	Plasma	DNA/BC	Other Biospecimens
단위: 명, 바이알(Unit: Participants, Vials)					
한국인유전체역학조사사업(Korean Genome and Epidemiology Study, KoGES)					
대규모코호트(Population-Based Cohort)	224,852	1,650,583	1,690,425	520,072	250,897
유전환경요인코호트(Gene-Environmental Model Cohort)	19,667	167,078	164,318	52,510	38,230
지역사회코호트(Community-Based Cohort)	12,013	237,314	194,458	62,985	192,185
기타 코호트(Other Cohort)	13,372	10,508	10,756	29,019	10,859
소계(subtotal)	269,904	2,074,484	2,059,957	673,566	492,181
질병예방연구(Disease Prevention Research)					
국가감염병면역도조사사업(National Infectious Disease Sero-surveillance Study)	2,139	4,749	-	-	-
국민건강영양조사(The Korea National Health&Nutrition Examination Study)	58,891	239,840	116,656	84,044	194
소계(subtotal)	60,830	244,589	116,656	84,044	194
면역·감염병연구(Immunology and Communicable Disease Research)					
브루셀라균매염색조사사업(National Brucellosis Survey)	7,439	32,890	-	-	-
에이즈코호트(AIDS Cohort)	1,281	-	41,733	-	36,803
홍역면역도조사사업(National Measles Sero-surveillance Study)	34,536	94,877	-	-	-
한국 HPV 코호트사업(Korea HPV Cohort Study)	485	2,540	4,342	-	-
소계(subtotal)	43,751	130,407	46,075	-	36,803

그림 13. 인체자원 수집 현황

지역사회기반코호트 국민건강영양조사사업

필수항목은 알록에 검색조건을 입력하여 인체자원을 조회할 수 있습니다. Code-book 다운로드

키워드/조건	항목	단위	검색 조건	연결 연산
일반사망	조사지역		<input type="checkbox"/> 안성 <input type="checkbox"/> 안산	AND
	성별		<input type="checkbox"/> 남자 <input type="checkbox"/> 여자	AND
	연령	세	<input type="checkbox"/> 31-40 <input type="checkbox"/> 41-50 <input type="checkbox"/> 51-60 <input type="checkbox"/> 61-70 <input type="checkbox"/> 71-80 <input type="checkbox"/> 81-90 <input type="checkbox"/> 91-100 직접입력: [] - []	AND
과거 고혈압 진단 경혈유무	과거 고혈압 진단 경혈유무		<input type="checkbox"/> 예 <input type="checkbox"/> 아니오	AND
	↳ (상세검색) 기반조사 이후 진단 경혈유무		<input type="checkbox"/> 예 <input type="checkbox"/> 아니오	AND
	과거 당뇨병 진단 경혈유무		<input type="checkbox"/> 예 <input type="checkbox"/> 아니오	AND
	↳ (상세검색) 기반조사 이후 진단 경혈유무		<input type="checkbox"/> 예 <input type="checkbox"/> 아니오	AND
	과거 위염/위궤양 진단 경혈유무		<input type="checkbox"/> 예 <input type="checkbox"/> 아니오	AND
	↳ (상세검색) 기반조사 이후 진단 경혈유무		<input type="checkbox"/> 예 <input type="checkbox"/> 아니오	AND

> 조회

그림 14. 인적 자원 검색

이렇게 수집된 방대한 양의 인체자원은 한국인의 질병 원인 규명, 진단 및 치료법 연구·개발에 활용되고 있으며 이를 통해 향후 개인 맞춤 치료법 개발 및 치료법에 의한 부작용 발생률의 완화를 기대할 수 있다.

IV. 결론

빅데이터는 현재 신가치 창출의 핵심으로 각 국가와 기업 들은 행정, 의료, 복지, 교육 등의 공공 및 민간 분야에서 이미 다양하게 활용되고 있으며 이를 통해 수많은 빅데이터가 저장·관리되고 있다.

이에 본 고에서는 그중 의료분야의 실제 빅데이터 활용 사례를 통해 의료분야와 빅데이터 융합의 효율성을 살펴보고 있으며 이러한 빅데이터 융합은 다음과 같은 효율성을 제공한다.

- 빅데이터 분석 기술의 발전으로 질병 및 신체 데이터의 축적을 통해 질병에 대한 사전방지 또는 대책을 마련함으로써 질병에 걸릴 확률을 낮출 수 있으며 이에 따른 의료비의 지출을 막을 수 있다.
- 빅데이터를 통해 의료분야의 방대한 양의 데이터를 사용자가 필요한 유의미한 데이터로 가공하여 제공함으로써 사용자는 자신이 원하는 의료 정보를 제공 받을 수 있다.
- 빅데이터를 통해 의료분야의 데이터 활용 범위가 증가 하면서 다양한 개인 건강기록을 수집·관리하고 분석함으로써 자신에 가장 적절하고 효율적인 치료법이나 의 약품을 찾는 등의 개인 맞춤형 의료서비스가 가능하다.

이와 같이 의료분야에서의 빅데이터의 활용은 기존의 의료 서비스보다 효율적이고 지능화된 의료서비스의 제공을 가능하게 하며 질병 방지 및 최적의 치료법의 제공을 통해 불필요한 진료비의 부담을 줄여줌으로써 오늘날의 다양한 스마트 디바이스를 활용한 의료서비스는 물론 향후 맞춤형 의료서비스의 제공을 위한 가장 효율적인 방안이라고 할 수 있다.

참고문헌

- [1] McKinsey Global Institute, “Big Data: The Next Frontier for Innovation”, for Innovation, Competition, and Productivity, McKinsey Inc
- [2] 김미혜, 김소희, “만성질환 노인의 의료비부담 관련요인에 관한 연구”, 한국사회복지학, 제48권, pp. 150~178, 2002

- [3] 보건복지가족부, “2008년 노인실태 조사 - 전국 노인생활실태 및 복지욕구조사”, 기초분석보고서, 2009
- [4] 양정선, “노인 가계의 의료비 지출과 부담에 관한 연구”, 한국가정관리학회지, 제 25권, 제 1호, 2007
- [5] 송태민, 이상영, 이기호, 박대순, 진달래, 류시원, 장상현, “U-Health 현황과 정책과제”, 한국보건사회연구원, 2011
- [6] 김상락, 강만모, 박상무, “빅데이터가 여는 미래의 세상”, 산업조직연구 6월호, pp.18~24, 2012
- [7] 안창원, 황승구, “빅데이터 기술과 주요 이슈”, 산업조직연구 6월호, pp.10~17, 2012
- [8] 유지연, “세계경제포럼(WEF)을 통해 본 빅데이터 논의 동향과 함의”, 방송통신정책, 제24권, 4호, 2012
- [9] 송태민, “보건복지 빅데이터의 효율적인 활용방안”, 보건복지포럼, 제193호, 2012, 11
- [10] 윤미영, 권정은, “빅데이터로 진화하는 세상-빅데이터 클러벌 선진사례”, 한국정보화진흥원 빅데이터 전략연구 센터, 2012
- [11] 이성훈, 이동우, “빅데이터의 국내외 활용 고찰 및 시사점”, 디지털정책연구, 제11권, 제2호, pp.229~233, 2012, 02
- [12] 정지선, “신가치창출 엔진, 빅데이터의 새로운 가능성과 대응 전략”, IT & Future Strategy, 제 18호, 2011
- [13] 채승병, 안신현, 전상인, “빅데이터: 산업 지각변동의 진원”, CEO Information, 851호, 2012
- [14] 행정안전부, 방송통신위원회, 지식경제부, “빅데이터 시대의 국가발전전략”, 2012 국가정보화백서, pp. 49~58
- [15] Gartner, www.gartner.com/newsroom/id/1731916
- [16] Pillbox, pillbox.nlm.nih.gov/
- [17] 2013 고령자 통계, http://kostat.go.kr/portal/korea/kor_nw/2/1/index.board?bmode=read&aSeq=260368
- [18] 구글 독감 예보, www.google.org/flutrends/
- [19] 구글 검색 트렌드, www.google.org/trends/
- [20] 유전자정보제공, www.1000genomes.org/
- [21] 한국인체자원은행, kbn.cdc.go.kr/

저 자 소 개



김 관 형
 2002: 한국해양대학교
 전자통신공학과 공학박사
 현 재: 동명대학교
 컴퓨터공학과 교수
 관심분야: 지능제어, 신호처리,
 헬스케어 시스템,
 의료정보 시스템



이 준 연
 1993: 중앙대학교
 컴퓨터공학과 공학박사
 현 재: 동명대학교
 미디어공학과 교수
 관심분야: 운영체제, 네트워크,
 모바일 프로그래밍,
 헬스케어 시스템



오 암 석
 1984: 부산대학교
 전자계산학과 이학사.
 1986: 중앙대학교
 컴퓨터공학과 공학석사.
 1997: 부산대학교
 컴퓨터공학과 공학박사
 현 재: 동명대학교
 정보공학부 교수
 관심분야: 데이터베이스,
 헬스케어,
 의료정보시스템,
 빅 데이터