

토론회

**심평원 및 약학정보원
개인질병정보 판매 행위로 본
현 정부의 보건 의료 빅데이터
추진 전략의 문제점**

일시 | 2017년 11월 27일(월) 13시30분

장소 | 국회의원회관 제8간담회실

주최 | 김상희 의원, 남인순 의원, 정춘숙 의원, 윤소하 의원, 참여연대, 건강과대안, 인도주의실천의사협의회, 진보네트워크센터, 무상의료운동본부, 경제정의실천시민연합, 건강세상네트워크

프로그램

13:30 사전행사 인사말

13:40 사회 박성용 한양여대 경영과 교수, 경실련 소비자정의센터 운영위원장

13:40발표1 심평원 사건을 통해 본 보건의료 빅데이터 추진 전략 문제점

정형준 인도주의실천의사협의회 정책국장, 의사

14:00발표2 개인정보 비식별화의 문제점

이은우 정보인권연구소 이사, 법무법인 지향 변호사

14:20 토론1 변혜진 건강과대안 상임연구원

이찬진 참여연대 사회복지위원회 실행위원, 변호사

김병수 성공회대학교 열림교양대학 교수

김록영 건강보험심사평가원 빅데이터부 부연구원

오상윤 보건복지부 의료정보정책과 과장

정영수 행정안전부 개인정보보호정책과 사무관

15:10 종합토론

목차

인사말	김상희 국회의원	04
인사말	남인순 국회의원	06
인사말	정춘숙 국회의원	09
인사말	윤소하 국회의원	11
발제1	심평원 사건을 통해 본 보건의료 빅데이터 추진 전략 문제점 / 정형준	13
발제2	개인정보 비식별화의 문제점 / 이은우	32
토론1	변혜진	59
토론2	이찬진	71
토론3	김병수	75
토론4	김록영	79
토론5	오상윤	80
토론6	정영수	81

인사말

김상희 / 더불어민주당 경기 부천시 소사구



안녕하세요? 더불어민주당 국회 보건복지위원회 김상희 의원입니다.

오늘 「현 정부 ‘보건의료 빅데이터’ 추진 전략의 문제점」 토론회에 참석해주신 모든 분께 감사인사 드립니다.

OECD 보건·의료 데이터 거버넌스 수준통계(2015)에 의하면 우리나라의 보건·의료 빅데이터 수준이 OECD 22개국 중 2위로 평가되고 있습니다.

건강검진 의무화를 통해 축적된 건강데이터와 양질의 의료데이터를 갖추고 있기 때문이라는 분석입니다. 1위인 아이슬란드의 인구가 33만 명인 것을 감안하면, 인구 5천만 명이 넘는 데이터를 관리하는 우리나라의 수준은 세계 최고라 할 수 있습니다. 하지만 국내의 건강·의료를 통합한 빅데이터의 활용수준은 상대적으로 미미합니다.

국민건강보험 제도와 우수한 병원정보 인프라 등을 통해 제각각 축적된 보건·의료 데이터를 통합해 우리나라 공공의료시스템을 강화하는 것은 꼭 필요한 방향입니다. 이를 통해 질병의 체계적인 예측·감시 시스템을 구축해 메르스 사태와 같은 상황을 예측하는 데 활용하고, 건강정보 통합서비스 등을 만들어 어느 병원에서도든지 쉽게 진료정보를 공유해 국민이 안전하게 치료를 받을 수 있도록 해야 합니다. 또한, 의학·정책 연구에 활용하여 소외계층 등 의료사각지대를 최소화할 수 있는 보건·의료정책을

개발에도 활용할 수 있을 것입니다.

이렇듯 ‘보건의료 빅데이터 사업’은 우리나라 보건·의료 정책과 의료산업 발전에 꼭 필요한 사업임에도 불구하고, 그동안 정부의 허술한 빅데이터 관리 실태로 인해 추진하는 데 애를 먹고 있습니다. 특히 건강보험심사평가원이 2014년 7월부터 2017년 8월까지 민간보험사 등에 6,420만 명의 건강정보 등을 제공한 사건과 약학정보원이 2011년부터 2014년까지 미국의 IMS헬스라는 빅데이터 기업에 국민 약 4천만 명, 약 50억 건의 처방전 정보를 판매했다는 사건은 빅데이터를 관리하는 정부의 신뢰성을 떨어트렸습니다.

오늘 「현 정부 ‘보건의료 빅데이터’ 추진 전략의 문제점」 토론회는 그동안 정부의 빅데이터 관리 실태를 돌아보고, 현재 추진 중인 ‘보건의료 빅데이터’ 사업계획의 문제점을 돌아보고자 합니다. 보건의료 빅데이터 사업을 추진하는 과정과 앞으로 구축될 보건의료 빅데이터 플랫폼 운영에 있어서 신뢰성 확보와 책임성 강화방안에 대해서 다양한 토론이 이루어질 것으로 기대합니다.

토론회에 참석하여 발표해주실 발제자·토론자분들, 이 문제에 대해서 함께 고민해 주시는 시민단체 그리고 ‘보건의료 빅데이터 플랫폼’을 구축하고 운영할 정부 각 부처의 공무원 여러분 감사드립니다. 오늘 토론을 통해서 국민에게 신뢰받은 보건의료 빅데이터 플랫폼을 구축해 우리나라의 공공의료와 의료서비스가 한 단계 도약할 수 있는 시간이 되길 소망합니다.

감사합니다.

인사말

남인순 / 더불어민주당 서울 송파구병



반갑습니다. 국회 여성가족위원장과 보건복지위원으로 일하고 있는 더불어민주당 남인순 국회의원입니다.

풍성한 결실과 수확의 기쁨을 나누는 만추의 계절에 ‘현 정부 보건의료 빅데이터 추진전략의 문제점에 대한 토론회’를, 존경하는 김상희 의원님과 정춘숙 의원님, 윤소하 의원님, 그리고 건강과대안, 인도주의실천의사협의회, 참여연대, 진보네트워크, 무상의료운동본부, 경제정의실천연합, 건강세상네트워크 등 시민사회단체들과 함께 열게 되어 무척 기쁘고 뜻깊게 생각하며, 관심을 갖고 자리에 함께하여 주신 모든 분들께 감사드립니다. 토론회를 준비하느라 애써주신 각 의원님실과 시민사회단체 관계자 여러분의 노고에 따뜻한 감사와 격려의 마음을 전하며, 바쁘신 가운데에도 좌장을 맡아 토론회를 진행하여 주실 경실련 소비자정의센터 운영위원장이신 박성용 한양여대 교수님과 발제를 준비해 주신 인의협이 정형준 정책국장님, 정보인권연구소 이사이신 이은우 변호사님, 그리고 참석해주신 패널 여러분들께도 감사를 드립니다.

5.9 촛불대선으로 출범한 문재인 정부는 100대 국정과제와 별도로 ‘불평등 완화와 소득 주도 성장을 위한 일자리경제’, ‘교육·복지·노동 체계 혁신으로 인구절벽 해소’,

‘국가의 고른 발전을 위한 자치분권과 균형발전’과 함께 ‘4차 산업혁명을 선도하는 혁신 창업국가’를 4대 복합·혁신과제로 선정하였습니다.

4차 산업혁명이 초래하는 산업구조·소득분배·생활방식 등의 변화과정을 포용적 성장을 위한 국가적 방향 전환의 계기로 삼고, 과학·기술혁신, 전 산업의 지능화, 제도개혁, 교육·공공·사회혁신 등을 통해 새로운 산업과 일자리가 창출되는 혁신 선도국가로 도약한다는 목표를 제시하였습니다.

이를 위해 대통령 직속 ‘4차 산업혁명위원회’를 신설하고, 인공지능 소프트웨어, 하드웨어, 데이터·네트워크 각 분야별 핵심 원천기술 및 이를 활용한 융합기술 개발 등 과학·기술혁신으로 초지능·초연결 기반을 구축하며, ICT 신기술·서비스 등 신산업 성장을 위한 제도를 정비하고, 고령층과 장애인 등 취약계층을 돕는 기술을 개발한다는 계획입니다.

하루가 다르게 발전하는 신기술의 홍수 속에서, 새로운 사회를 맞을 준비를 해야 한다는 목소리와 소외되는 시민들을 보호해야 한다는 목소리가 함께 높아지고 있습니다. 새로 도입되는 기술을 손쉽게 접하기 어려운 다수의 국민들은, 소위 ‘4차 산업혁명’이라는 표현이 반갑지만은 않을 것이라고 생각합니다. 특히 일자리 문제, 의료·건강 문제, 개인정보 문제 등에 있어서는 단순히 기술만을 생각할 것이 아니라 개인의 권리와 보호가 매우 중요하게 고려되어야 합니다. 이러한 가운데, 소중한 건강정보의 무분별한 사용을 막고, 이를 안전하게 관리하기 위한 법적, 제도적 기틀이 논의될 이 자리가 더없이 귀중하게 느껴집니다.

미국, 유럽, 호주, 그리고 특히 최근 이슈가 된 영국 등에서는 보건의료 빅데이터를 활용하는 데에 있어 사회적인 공감대가 형성되어 있는 것으로 알고 있습니다. 의료기술개발, 공공정책·의료전달체계 개선, 치료법간 효과비교 등 철저히 공공적 목적으로, 연구 중심으로 활용되고 있습니다. 또한 이러한 방향을 결정하는 데 있어, 정부의 정책방향만이 아니라 시민사회 및 의학전문가, 법조계·정보보호 전문가 등의 의견이 크게 반영되고 있습니다.

그러나 최근 알려진 우리나라의 보건의료 빅데이터 관련 정책의 내용과 그 추진 과정은 이러한 국제적인 추세와 크게 다른 것이 아닌가 우려되고 있습니다. 처음 논의를 시작할 때부터 산업적 활용을 염두에 두고 검토하고, 전략 수립 과정에서 이를 우려하는 시민사회의 목소리는 배제되었다는 지적이 제기되었습니다. 또한 지난 국정감사 과정에서 정춘숙 의원께서 지적한 바와 같이, 건강보험심사평가원이 국민 건강정보를 민간 보험사에게 제공해 온 사실이 드러났습니다. 이러한 부분은 시정되어야 할 것입니다. 건강정보의 주인인 국민의 이익에 반해서 상업적으로 사용할 우려

가 있는 민간 기업은 활용 주체에서 배제되어야 할 것입니다.

한편, 보건의료 빅데이터는 잘 활용된다면 여성·장애인·저소득층 등 소외계층의 건강 실태, 소득·지역에 따른 건강 불평등, 지역별 산업재해나 암 발생률 등 우리 국민들의 건강 상황을 한 눈에 볼 수 있는 좋은 창이 될 것이라고 생각합니다. 과거의 낡은 의사결정 구조와, 소통 방식을 고수하고 정책을 추진할 경우 그 어떠한 공감도 얻기 어렵습니다.

최근 국회 보건복지위원회의 2018년도 보건복지부 예산안 심의과정에서 제가 강조한 바 있듯이 보건의료 빅데이터는 상업적 활용을 배제하고 철저하게 공공적 목적에 한정하여 활용하되, 보건의료 정보의 원천적 권리 소유자인 개인을 특정할 수 없도록 주민번호와 생년월일 등을 모두 삭제하여 개인정보를 철저히 보호하여야 하고, 시민사회를 포함한 각계각층의 의견을 모을 수 있는 위원회를 구성해야 합니다. 이렇게 구성된 위원회는 단순한 자문 역할에 그칠 것이 아니라, 보건의료 빅데이터 정책을 추진함에 있어서 실질적 심의 기구로서 역할을 하여야 할 것입니다.

저는 이러한 제도가 법적인 근거를 가지고 안정적으로 추진될 수 있도록 하고, 건강정보를 국민의 건강증진과 공익을 위해 활용하며, 이를 오·남용하거나 허술하게 관리하는 이들을 처벌할 수 있도록 법제화를 추진하는 데 국회차원에서 노력하겠습니다.

이러한 점에서 오늘‘현 정부 보건의료 빅데이터 추진전략의 문제점’에 대한 토론회는 의미가 크다고 생각합니다. 보다 깊이 있는 토론을 통하여 정부의 보건의료 빅데이터 관련 정책을 점검하고 부족한 사항들을 시정·보완하여 4차산업혁명 시대 바람직한 보건의료 빅데이터의 통합과 분석, 개인정보 보호 및 공공적 활용방안을 모색하는 소중한 계기가 되길 바랍니다. 오늘 논의된 사항에 대해서는 국정에 적극 반영하여 주실 것을 당부드리며, 저 또한 국회 정책 및 입법활동에 반영하도록 하겠습니다. 참여해 주신 모든 분들께 다시 한 번 감사드리며, 건강과 행복, 평화를 기원합니다.

인사말

정춘숙 / 더불어민주당 비례대표



안녕하십니까? 국회 보건복지위원회 더불어민주당 정춘숙 의원입니다.

먼저 추운 날씨에도 불구하고 우리나라 보건의료 빅데이터에 대해 많은 관심을 갖고 토론회에 참석해주신 여러분에게 큰 감사의 인사를 드립니다. 아울러 토론회의 발제와 토론을 맡아주신 모든 분들께도 진심으로 감사의 인사를 드립니다.

제가 이번 국정감사 당시 밝혔듯이, 건강보험심사평가원은 지난 3년 동안 삼성생명, KB생명 등 민간보험사와 보험연구기관에 건당 30만 원을 받고 약 1억 명이 넘는 진료데이터 87건을 제공한 것으로 드러났습니다.

심평원이 제공한 '표본데이터셋'은 개인건강정보에서 이름, 주민번호 등 개인정보는 제외한 데이터였지만, 성별 연령 등을 포함해 진료행위 상병내역 주상병 등이 담긴 진료내역과 원외처방내역으로 구성된 중요한 자료였습니다.

심지어 민간보험사가 보험 위험률 산출, 보험상품연구와 개발 등과 같은 영리목적으로 활용하겠다고 밝혔음에도 심평원은 표본데이터셋을 제공할 때 "학술연구용 이외의 정책, 영리목적으로 사용불가"하다는 서약서만 받고 제공한 사례도 있었습니다.

반면 심평원과 유사한 국민건강 빅데이터를 보유한 건강보험공단도 민간보험사에 대

해 '공공데이터 이용과 제공은 공익 목적으로 제3자의 권리를 침해하지 않아야 한다'며, '민간보험사의 경우 건강보험 진료데이터를 분석해 특정 질환 유병자, 기왕력자 또는 위험요인 보유자에 대해 민간보험의 가입차별 등에 악용될 가능성이 있어 국민 건강권 및 권리보호차원에 제공하지 않는다'는 이유를 들며 자료 제공을 하지 않았습
니다. 이것이 우리나라에서 보건의료와 관련하여 가장 거대한 빅데이터를 가지고 있
는 두 기관의 현실입니다.

이번 국정감사에서 저는 심평원이 민간에 제공한 빅데이터가 아무리 개인정보가 식
별되지 않는 자료라 하더라도 민간보험사에 제공될 경우 보험사의 보험상품개발과
민간보험 가입차별 등에 악용될 가능성이 충분히 있음을 지적하고, 제공을 중단할 것
을 보건복지부 장관에게 촉구하여 답변까지 받았습니다.

이 뿐 아니라 건보공단과 복지부, 심평원이 더불어 건강보험 정보의 공익성을 지키고
제3자의 권리가 침해되지 않는 범위 내에서의 빅데이터 활용기준을 만들어야 한다고
지적했습니다.

4차 산업시대에 보건의료 빅데이터가 중요한 만큼 그 활용방법과 공익성 그리고 권
리침해 부분도 심도 깊게 다뤄져야 합니다. 그리고 오늘 이 자리가 그 첫 걸음이 되
는 자리가 될 것이라고 기대합니다.

아무쪼록 오늘 토론회를 통해서 제시되는 다양한 의견과 논의가 우리나라의 보건의
료 빅데이터를 보다 건전하게 발전시킬 수 있는 자리가 되기를 바랍니다.

다시 한번 오늘 이 자리에 참석해주신 모든 분들에게 감사드리며, 건강과 행복이 가
득하시길 기원합니다.

감사합니다.

인사말

윤소하 / 정의당 비례대표



안녕하십니까, 정의당 국회의원 윤소하입니다.

먼저 「보건의료 빅데이터 추진 전략의 문제점」 토론회를 함께 준비해 주신 더불어민주당 김상희 의원님, 남인순 의원님, 정춘숙 의원님을 비롯해 참여연대, 건강과대안, 인도주의실천의사협의회, 진보네트워크센터, 무상의료운동본부, 경제정의실천시민연합, 건강세상네트워크 등 시민사회단체 회원 여러분께 존경과 감사의 인사드립니다.

지난 국정감사를 통해 건강보험심사평가원이 국민의 개인 건강정보를 민간보험사에 판매한 사실이 크게 이슈화 되었습니다. 또 약학정보원은 지난 2011년부터 2014년까지 미국의 빅데이터 업체 IMS헬스에 우리나라 국민의 약 50억 건의 처방전 정보를 팔았다는 것에 국민 모두가 놀랐습니다.

현재 대한민국 국민 개인의 정보는 개인정보보호법에 의거해서 보호를 받고 있습니다. 본인의 동의가 이루어지지 않은 상황에서는 개인의 정보가 그 어디에도 제공될 수 없습니다. 더욱이 개인의료정보는 개인정보보호법에도 민감정보로 구분하여 그 유출을 엄격히 관리하고 있음에도 불구하고, 이와 같은 일이 현실에서는 버젓이 벌어지고 있는 것이 대한민국의 개인정보 관리 현실입니다.

지난 박근혜 정부는 의료분야 규제완화와 서비스활성화를 이유로 의료정보를 담은 클라우드시스템활용 등을 추진했습니다. ‘맞춤형의료’, ‘정밀의료’ 라는 이름으로 의료기관에서만 생성, 관리될 수 있는 개인 의료건강정보를 민간기업에서도 생성관리 할 수 있도록 추진했습니다.

많은 시민사회단체와 국민들의 문제제기에도 불구하고 지난 정부가 추진하려던 의료정보 정책을 현 정부에서도 재추진하려는 움직임이 있습니다. 현재 보건의료 빅데이터 플랫폼구축이라는 이름으로 4개의 공공기관이 보유하고 있는 국민 개인 의료·건강 정보를 수집·연계할 수 있는 플랫폼을 구축하기 위한 사업을 추진 중이며, 내년도 사업예산이 편성된 상황입니다.

앞서 확인한 예처럼 국민 개개인의 의료·건강 정보가 민간 보험사를 비롯한 민간영역으로 흘러들어갈 수 있다는 우려는 여전히 존재합니다. 또한 해당 사업이 국민 개개인의 ‘어떤 정보’를, ‘어떤 목적’으로, ‘어떤 원칙’에 맞게 모으고, 사용할 것인지 구체적인 기준도 마련되어 있지 않은 상황이라는 점에서 시민사회를 중심으로 제기하는 여러 우려들이 기우만은 아니라는 점도 인정해야할 지점입니다.

오늘 토론회를 통해 정부는 사업을 추진하기에 앞서 부족함이 없는지, 국민이 갖고 있는 우려를 해소하기 위한 노력이 더 필요한 상황은 아닌지 다시 한 번 돌아보는 계기가 되고, 동시에 정부의 정책 구상에 국민의견을 반영하는 좋은 기회가 되길 바랍니다.

마지막으로 정부의 좋은 정책 추진에는 반드시 그에 맞는 동력이 필요합니다. 무엇보다 강한 동력은 국민의 동의와 합의과정이라는 입장을 가지고 이번 토론회가 그 합의를 만들어가는 길목에서 유의미한 토론회가 되길 희망합니다. 다시 한 번 오늘 찾아주신 모든 분들의 안녕과 건강을 진심으로 기원합니다. 감사합니다.

심평원 사건을 통해 본 보건의료 빅데이터 추진 전략의 문제점

정형준 / 인도주의실천의사협의회 정책국장, 의사

1. 들어가며

- 지난 10월 24일 더불어민주당 정춘숙 의원실은 건강보험심사평가원(이하 심평원)이 민간보험기관 등에 6,420만 명의 건강정보 등이 포함된 데이터를 제공했다고 발표했다. 문제는 민간보험사가 영리목적(보험상품의 손해를 평가 등)으로 정보를 활용한다는 것을 알고 있었음에도 2014년 7월부터 올해 8월까지 비식별화된 정보를 지속적으로 제공한 것임.
- 또한 약학정보원은 지난 2011년부터 2014년까지 미국의 빅데이터 업체 IMS헬스에 우리나라 국민의 약 4천만 명, 약 50억 건의 처방전 정보를 팔았음. 그리고 IMS는 우리나라 국민의 정보를 전 세계에 되팔고 있음. 관련하여 재판이 진행 중에 있으며, 이 사건은 우리나라 건강정보 보호의 허술함을 단적으로 보여주고 있음.
- 현재 우리나라는 건강정보를 포함한 개인정보가 빈번하게 유출되고 있어 대안 마련이 시급한 상황임. 그러나 정부는 이에 대한 대안 마련 없이 상업적 활용 목적을 골자로 한 보건의료 빅데이터 사업을 비공개로 추진하고 있으며, 이 사업에 대한 타당성 등에 대해 충분한 논의를 거치지 못하고 있음. 더욱이 추진 과정도 비공개로 진행되고 있어 정책추진 투명성이 의심됨.

- 특히, 개인건강정보와 관련한 활용, 적용, 결합등을 논의할 법률도 없는 상황에서 박근혜 정부 말기 막무가내도 도입된 ‘가이드라인’을 근거로 사용하는 상식밖에 현상마저 벌어지고 있음. 이에 지금까지 한국에서 추진된 빅데이터, 정밀의료, 개인건강정보 유출, 조작(비식별화), 결합 등의 문제점을 살펴보고 대안을 제시하고자 함이 이 글의 목적임

2. 의료산업화와 개인건강정보

- 직접적 의료민영화 사안인 영리병원, 부대사업확대, 영리자회사 설립 등은 박근혜 정부의 의료산업화 핵심 계획이었음. 그러나 국민여론과 시민사회단체의 역풍을 맞고 박근혜 정부 후반기부터는 앞서 밝히 직접 의료민영화 사안이 숨겨지고, 다른 의료산업화 사안들이 재배치되었음. 이를 본인은 ‘연성’ 의료산업화로 파악함. 대표적인 연성 의료산업화 과제로 재생의료(줄기세포, 유전자치료제 등)와 정밀의료(빅데이터, IT-의료연계 등)가 있었음.
- IT-의료연계는 2006년 ‘의료산업선진화위원회’때부터 주창되던 의료산업화론자들의 레퍼토리로 지난 10년 동안은 지난했던 ‘원격의료’ 논란을 중심으로 도입논쟁이 벌어졌음. 원격의료는 그 자체의 산업적 확대 측면 뿐 아니라, 민간보험사의 영리적 건강관리서비스 도입, 개인건강정보 수집, 축적의 선제적 과제이기도 했음. 그러나 ‘원격의료’는 의료공급자의 반발과 시민사회단체의 저항 등으로 쉽게 관철되지 못했음. 때문에 원격医료를 중심으로 거론되던 e-health 주장은 소강시점에 접어들게 되었음.
- 이런 시점에 정밀医료를 시발점으로 빅데이터 구축이 논의되는 것(그림1-1 참고)이 현 시기의 e-health의 특징임. 이는 IT-의료 연계의 ‘최신티렌드’라기 보다는 한국에서 e-health를 도입하려는 세력(IT 기업, 민간의료보험, 대형병원 등)이 외피를 바꿨다고 보는 것이 온당한 측면이 있음.
- 작년 촛불항쟁을 중심으로 박근혜 정부의 적폐청산, 그 중에서도 보건의료 적폐로 줄기세포 등이 거론되었던 상황에서 입증된 표적항암제 등을 중심으로 최근 벌어지는 보건의료 빅데이터사용, 유전자정보이용, 인공지능진단프로그램 도입 등의 논의는 확장성이 높음. 또한 보건의료가 여전히 새로운 차세대 산업먹거리로 제시되고 있는 측면에서 볼 때, 생명공학과 첨단의료기기 산업이 오래된 인프라가 필요한 측면에 비해, 축적된 의료정보를 기반으로 하는 ‘빅데이터’ ‘의료정보화’는 손쉬운 산업화로 포장되고 있음.

④ 4차 산업혁명 대비 성장기반 조성

- 보건의료빅데이터 플랫폼 구축(115억원, 신규), 바이오헬스 기술비즈니스 생태계 조성 지원 확대(30→36억원)
- 국가전략프로젝트(35→141억원), 국가 항암신약개발(76→146억원), 의료기기기술 개발(220→291억원), 라이프케어융합서비스개발(114억원, 신규) 등 연구개발 지원 확대
- 한약현대화, 한약 공공인프라 구축 등 한의학 산업 육성 강화(138→217억원)

4 4차 산업혁명 대비 성장기반 조성

< 주요 사업 편성현황 >

- 보건의료 빅데이터 플랫폼 구축 : ('17) 0→('18안) 115억원 (115억, 신규)
- 바이오헬스기술비즈니스생태계조성 : ('17) 30→('18안) 36억원 (6억, 20.0%)
- 의료기기산업경쟁력강화 : ('17) 42→('18안) 46억원 (5억, 11.0%)
- 한의학산업육성 : ('17) 138→('18안) 217억원 (79억, 57.0%)

□ 【보건산업】 지속가능한 성장을 위해 보건산업의 성과확산을 추진

- (보건의료 빅데이터 플랫폼 구축, 신규) 공공기관 보유 데이터 연계 시스템(77억, 기관 간 분석자료 공유·활용 서비스(24억) 등 신규 반영
- (바이오헬스 기술비즈니스 생태계 조성) (가칭)바이오헬스 기술비즈니스 종합지원단 설치·운영을 통한 우수기술발굴·가치향상·시장진출지원의 전주기 사업화 지원(6억, 신규)
- (의료기기산업 경쟁력강화) 마이크로 의료로봇 산업 성장 생태계 조성지원(3억, 신규) 및 해외 ODA 지원센터 구축(3.5억, 신규)
- (한의학 산업 육성) 한약 현대화(탕약표준조제시설 62억) 및 한약 공공인프라 (한약 비임상연구시설(GLP) 55억, 임상시험용 한약제제생산시설(GMP) 45억) 구축

<그림1-1> 보건복지예산안, 2017년 8월 29일 발표자료 중

- 따라서 현재 추진 중인 '빅데이터', '정밀의료'는 국제적으로 최근 각광을 받는 신산업의 측면보다는 기존의 의료산업화의 선결조건인 '의료정보화'의 산업적, 영리적 이용의 측면에서 바라볼 때 한국에서의 그 역사적 접근이 가능함. 또한 해외의 '빅데이터'가 주로 의료정보를 기반으로 한 진단프로토콜, 타겟약제개발 등이 주된 목적이었다면, 한국의 경우는 여전히 민간의료보험의 시장확장, 마케팅산업 등의 효율화의 도구로 사용되는 것이 주된 용도라는 점도 특이점임.
- 이런 맥락에서 박근혜 정부 기간 건강보험심사평가원이 민영보험의 손해율 등 경영효율화를 위해 공보험의 개인건강정보를 세트로 판매한 것을 역사적으로 이해

할 필요가 있음. 심사평가원이 거둔 수익은 고작 1,800만 원가량이지만, 민간보험이 거둔 이익은 수천억에 다를 수 있는데, 이것이 과연 금융산업의 효율화이지 의료산업화 효과라고 부를 수 있을지도 반문할 필요가 있어 보임.

3. e-health 도입수준과 그 기반 논의

- 의료산업화, IT기업 연계는 아날로그 방식으로 수행되는 의료를 디지털화 하는데서 출발 할 수밖에 없음. 이는 1) 높은 효율성을 추구하는 경향 2) 기존 아날로그 정보의 디지털화 3) 표준화 등으로 나타났고, 실제 병의원과 진료현장에서는 전자의무기록(EMR), 처방전달시스템, 디지털영상정보전달시스템(PACS)의 개발과 정착으로 드러났음. 이런 기반하에 디지털화된 의료를 보다 큰 시장으로 만들기 위해서는 정보표준화를 통한 호환 가능성을 높이는 것이 필요해졌고 이것이 보건 의료정보의 플랫폼 사업이 되었음.
- 국내에서도 이를 위한 의료 용어 표준화, 전자의무기록, 처방전달시스템, 디지털 영상정보 전달시스템 등의 표준화 시도가 진행되었으나, 병의원 기관마다 각자 다른 시스템을 갖고 있어 난제가 된 바 있음. 특히 한국의 경우 의료공급이 전적으로 민간주도여서 상호 경쟁적인 표준이 난립하여 표준화는 공적의료제도를 가지고 있는 나라에 비해 더 어려운 과제가 되었음. 이는 시장화 되어있는 의료공급으로 인해 의료정보산업에서는 뒤쳐지지는 아이러니를 낳게 된 셈임.
- 결국 이런 시스템을 하나로 표준화하는데 있어 드는 비용문제가 계속 제기되고 있으며, 대자본(대기업) 중심으로의 표준화되는 것에 대한 중소병의원 등의 갈등도 초래되었음. 과도한 공급시장화로 표준화의 난제가 계속 발생하고 있으며, SK 조차 의원급 진료프로그램인 ‘의사랑’등에 진출했다가 지분을 매각하고 철수한 상황임.
- 결국 2014년이 되어서 정보 보안에 대한 대중의 우려를 불식시키겠다는 명목으로 의료/건강 정보 보안 산업 활성화가 주창되었음. 2016년 의료법 개정으로 의료기관 의료정보 보안 업무의 외주화 허용, 클라우드 시스템 활용 등이 추진하였음. 박근혜 정부는 2016년 초 클라우드 서비스 활성화를 천명했으나, 역시 법률이 아닌 행정독재식 방법으로 이를 처리한 바 있음.
- 박근혜 정부 말기에는 의료기관에서만 생성, 관리될 수 있는 의료, 건강 정보를 민간기업도 생성, 관리할 수 있도록 시도했음. 이는 앞서 말했듯이 원격의료, 건강관리서비스, 의료기관 외 유전자 검사 서비스 등의 전제조건이었음.

- 특히 기존 허가된 유전자 검사로는 키, 몸무게, 혈액형 등 10개 정도의 크게 유용한 정보 검사만이 제공되었지만, 최근DNA검사 기계 등의 발전으로 단시간 내에 수백가지의 개인 정보가 산출된다는 점에서 이런 유전자정보의 민간기업 취급 확대허가는 충분한 사회적 검토가 필요한 것이었음. 물론 정부는 이러한 검사를 ‘맞춤형의료’ ‘정밀의료’ 라는 이름으로 허용하려했음.
- 이런 개인건강정보와 의료정보가 결합될 때의 문제, 그리고 이런 민간기업의 개인건강정보 수집과 유통이 나올 문제가 제기되는 와중에 2016년 6월에는 개인건강정보의 처리에 관해서도 단순한 가이드라인으로 이를 관철시키려 하였음(그림 1-2).



<그림1-2> 개인정보 비식별화 조치 가이드라인. 2016년 6월

- 다음으로 의료기관에서 생성되어 건강보험공단, 심평원 등에 보관되어 있는 대규모 의료/건강 정보를 제약 없이 활용할 수 있도록 하였음(그림1-3). 이는 ‘빅데이터 산업’으로 불리며 4차 산업혁명의 신산업으로 포장되었음. 이미 성상철 건강보험공단이사장은 2015년 신년사에서 빅데이터산업을 공단이 추진할 사업으로 직접 언급한 바 있음.

건강보험 빅데이터, 누구나 편리하게 이용할 수 있게 된다.

- 전국 16개소 빅데이터 분석센터 확대 운영 -
- 복지부·공단·심평원 건강보험 빅데이터 활용 협의제도 출범 -

- 앞으로 민간에서 연구나 사업화 목적으로 건강보험 빅데이터를 활용하려는 경우 보다 편리하게 접근하여 맞춤형으로 활용할 수 있도록 정부와 빅데이터 보유기관(건강보험공단, 건강보험심사평가원) 간 협의제를 통해 지원하고, 전국 16개소의 빅데이터 분석센터에서 필요한 데이터를 제공하여 분석·처리할 수 있는 분석공간을 제공한다.
- 보건복지부(장관 정진엽), 국민건강보험공단(이사장 성상철), 건강보험심사평가원(원장 손명세)은 9월부터 ‘(가칭)건강보험 빅데이터 활용 협의제’를 출범하고, 직접 방문하여 데이터 분석·처리가 가능한 빅데이터 분석센터 총 16개소를 본격 운영한다고 밝혔다.

<그림1-3> 2016년 8월 30일 보건복지부 보도자료

- 건강정보의 빅데이터 제공은 제약기업, 민간의료보험 회사, 건강관리서비스 회사 등의 마케팅(상품 판매 및 판촉)을 위한 전제조건이기도 함. 문제는 이런 논란에도 보건복지부가 작년 박근혜정부가 발표한 6월 30일 <개인정보 비식별 조치 가이드라인>에도 참여했다는 점임. 보건복지부는 건강정보와 관련된 국민건강보험의 우선권과 공적이용을 중심으로 이를 방어해야 하는 부서라는 점에서 이는 매우 우려스러운 행동이었음. 빅데이터는 행안부에서 사용된 내용도 이미 심각한 문제점이 드러난 상황임.

4. 보건의료 빅데이터 활용의 효과 및 전망은 충분히 입증되었나?

- 보건의료 부문 빅데이터 활용의 효과는 아직 충분히 검증되지 않았음. 치료의 질 및 효과의 향상, 질병 예방, 환자 안전 수준의 향상, 의료비 절감 등의 효과가 거론되며 최근 논의가 가속화 되는 것은 사실이지만, 기술적, 사회적으로 이는 아직 미완의 상태임. 많은 논의와 장밋빛 전망에 견줘 실제 현실에서 데이터로 입증된 효과를 보이는 보건의료 빅데이터 활용 모델은 매우 적음.
- 올해 정부가 2017년 9월 5일 발족한 ‘정밀의료 사업단’의 첫번째 사업은 2가지임. 첫번째 ‘암 정밀 진단치료법 개발 사업단’, 두번째 ‘정밀의료 병원정보시스템 개발 사업단’임. 두번째 ‘정밀의료 병원정보시스템 개발 사업단’은 앞서 설명한 의료정보 표준화 및 보안관리 를 위한 선결과제사업에 지나지 않음(그림1-4).



<그림1-4> 2017년 9월 5일 ‘정밀의료 사업단’의 2번째 개발사업

- 이는 엄밀히 정밀의료의 효과 및 안정성에 대한 논의가 아니라, 정밀医료를 위한 선결조건인 표준 데이터 축적을 위한 인프라 개발임. 이는 아직까지 쓸모있는 ‘빅데이터’가 제대로 축적되지 못하고 있다는 인식의 발로이며, 표준화된 ‘빅데이터 셋’으로 정밀의료의 가시적 효과를 보이는 것이 현재 수준에서는 요원하다는 뜻임(그림1-5).

- 이와 관련하여 **이상현 사업단장**은 “**재정여건이 열악한 지방·중소 병원의 낙후된 병원정보시스템을 대체하고, 보안 전담인력 부재에 따른 개인의료정보 유출 우려를 해소 하는 등 정밀의료 병원정보시스템의 활용도는 매우 높을 것**”이라는 견해를 밝혔다.

<그림1-5> 병원정보시스템의 표준화 및 보안상향을 통한 정보표준화가 목적임을 밝힘

- 첫 번째 과제사업으로 밝힌 ‘**암 정밀 진단치료법 개발 사업단**’의 경우도 빅데이터를 이용하는 사업이라고 보기에 어려움. 우선 표적치료제 적용이 가능한 약 2000명을 발굴하고 3건의 표적치료제 개발을 추진하는 임상시험 프로그램임. 이는 다양한 유전정보나 개인건강정보를 토대로 신약을 개발하는 ‘정밀의료’의 광고 효과와도 차이가 있음. 이미 표적항암제의 효과가 적용된 유전자변이에 대해 타 암종에서도 적용이 가능한지 살펴보는 것은 확장된 ‘임상시험’에 지나지 않음(그림1-6).

- 임상시험에는 **새로 개발되는 항암제뿐 아니라 이미 허가된 표적치료제의 적용 질환 확대 등 다양한 접근을 시도하여, 신약에 대한 우리나라 암환자의 접근성을 최대한 향상시키고자 한다.**

◇ 바구니형(Basket) 임상시험	: 다양한 암종에서 관찰되는 동일 유전자 변이에 대해 단일 항암제 처방가능성을 입증하기 위한 임상시험
◇ 우산형(Umbrella) 임상시험	: 다양한 유전자 변이에 의해 발생한 단일 암종에 대해 여러 가지 항암제를 처방하는 임상시험

<그림1-6> 허가된 항암제의 타 암종 및 동일 유전자에 적용되는지 확인하는 임상시험이 목적임을 밝힘

- 결국 이런 선결과제 및 기존 임상시험을 확장하는 사업을 통해 국민의료비가 절감되는 것으로 보도하는 행태는 문제임. 우선 특정유전자에 대한 표적항암제의 효과는 이미 과학적으로 입증이 되어 있으며, 현재 상당수 표적항암제의 대상이 되는 암환자가 유전자검사를 실시하고 있는 사안임. 결국 암이 발견되지 않은 전국민을 대상으로 특정 유전자조사를 하는 프로그램과 생활양태 등을 결합하는 빅데이터 사업이 아닌 마당에야 기존 표적항암제 임상시험과 이번 ‘빅데이터’로 포장된 사업은 전혀 다를 바가 없음.

- 특히 정보표준화사업은 더 많은 데이터를 표준화 하여 축적하면 무언가 ‘개인맞춤 치료’가 가능하지 않을까 하는 가정에서 출발하는데, 이런 가정을 확인하기 위한 기반사업에 무작정 수십억을 배정한 꼴임. 또한 정보표준화를 위해 향후 민간 중심 한국의료체계는 더욱 더 큰 장애가 될 것이 분명한 상황에서 경쟁력 있는 사업인지도 재검토가 요구됨.
- 역사적으로 1990년경부터 철저히 유전자 중심적 관점에서 출발한 인간게놈프로젝트도 당시 맞춤형 유전자로 질병치료의 새로운 장을 연다는 것이 목적이었음. 그러나 2003년에 이 프로젝트는 사실상 종료되었는데, 실제 유전자의 90% 이상을 분석하였으나, 이를 통해 밝혀진 질병이 0.1% 수준으로 사실 유전자 중심의학의 폐기를 앞당겼던 것이 아이러니 하게도 인간게놈프로젝트였음.
- 게놈프로젝트의 역설은 역으로 우리에게 유전자의 과학적, 사회적 의미를 되새겨 볼 것을 요구했음. 즉 역동적 생명현상을 유전자로만 환원할 것이 아니라 다양한 맥락에서 고려하는 것이 필요하다는 성찰이었음. ‘유전자’, ‘생명체’, ‘환경’의 능동적 상호관계를 ‘삼중나선(Triple helix)’으로 표현했던 생물학자 르윈틴의 주장을 진지하게 재검토하고, 식품, 운동, 환경 등의 요소에 대한 심층적 분석으로 의학적 성찰이 변화하는 계기를 제공했음.
- 그러나 최근 구미각국의 당시 유전자정보 등을 중심으로 벌어졌던 신기루현상을 타고 개인건강정보 빅데이터 사업이 재차 조명 받는 것은 충분한 논의와 성과가 없다면, ‘가정’에 지나지 않는 것임. 특히 이러한 유전자정보 및 건강정보를 이용한 사업은 대부분 민간이 주도하고 있음. 만약 이런 유전적 정보를 대규모 정보로 집적화하여 효용화 하려한다면 이는 기본적으로 공적영역에서 우선적으로 진출해야 할 것임.

5. 국민 개인의 건강정보를 배타적으로 보호하고, 건강정보를 매개로 한 감시, 차별, 배제, 낙인이 이루어지지 않도록 보호할 수 있는가?

- 비식별화된 개인건강정보의 사용이 가장 활성화 되어 있는 미국의 경우도 이미 1996년 건강보험 정보의 이전 및 그 책임에 관한 법률(Health Insurance Portability and Accountability Act, HIPPA)를 통해 의료관련 행정 및 금융자료의 전자교환을 표준화 하는 법률을 제정한 바 있음. HIPPA는 일상적인 의료업무 차원에서 수집되는 개인 건강정보를 그 이외의 목적으로 이용하는 것을 이차적 이용을 규정하여 원칙적으로 금지하고 있음.
- HIPPA에 따르면 개인건강정보를 표준화, 비식별화, 변환작업 등을 수행하는 것은 건강정보 교환소(Health Care Clearinghouses)의 역할임. 이는 별도의 법안(Health Care Clearinghouse Law)의 지배를 받으며, 건강정보를 관리하는 제3 공공기관인 셈임. 이곳에서 개인을 식별할 수 있는 모든 건강정보를 보호하고 있음. HIPPA, Health Care Clearinghouse를 통해 비식별화한 건강정보의 활용과 공개에는 제약이 없으나, 이를 위해 비식별화의 방법, 책임, 유통 등에 최소한의 규제가 존재하고 있음.
- 반면 한국의 경우 어떠한 개인건강정보를 관리, 유통, 조합할 수 있는 법률적 근거는 없음. 여기에 한국은 핵심 개인식별번호인 주민번호가 존재하는 나라임. 이는 유럽, 미국 등에서 진료 및 투약 등 핵심건강정보를 심사평가등을 위해 공적 보험에 제공할시 생년월일과 거주지역, 이름 등으로 조합(최소 3가지 정보가 조합되어야 개인식별이 용이)해서 개인식별을 하는 경우와 달리, 단독 번호만으로 정보를 식별하는 방식임을 말함. 즉 정보식별성이 매우 높다는 뜻임(최고수준임). 거기다 전 세계에 유례없는 신용카드 보급률, 금융실명제, 핸드폰 과 인터넷 보급, 택배거래 활성화 등은 개인정보유출의 식별화 가능성을 한층 더 높이고 있음.
- 이런 상황에서 빅데이터 주창자들은 주민번호 및 이름만 삭제한 비식별화 방식의 데이터셋을 광범하게 이용하자는 주장까지 서슴지 않고 꺼내는 상황임. 이러한 비식별화에 대한 별도의 대안이나 보완 방식에 대한 공유, 거버넌스 확보 등도 없고, 오로지 근거는 박근혜 정부의 가이드라인일 뿐임.

6. 박근혜 정부 비식별화 가이드라인의 문제점

- 비식별화 가이드라인은 그 내용에서도 수많은 문제점이 존재함. 그 첫째는 비식별화 평가를 해당기관의 개인정보 보호책임자가 구성하는 부분임. 그 조차 3명이상의 관련분야 전문가로 명시되어 있음.
- 여기에 정보집합물 결합을 위해 개인별로 부여된 식별자가 매칭키로 사용되는데, 이 매칭키가 ‘주민번호’를 대체하는 ‘임시 대체키’이나, 사실상 주민번호에 한차례 매칭된다는 문제점이 존속함. 또한 결합을 제공하는 기관은 단일 제3기관이 아니고, ‘산업 내 기업간 결합은 해당 분야 전문기관’으로 명시되어 한국인터넷진흥원, 한국정보화진흥원등이 추천단위임. 이들 단위가 비식별화된 개인건강정보를 제대로 결합해 문제없이 제공할 수 있는지는 차치하더라도, 결합기관의 난립이 가져올 문제점조차 가이드라인은 무시하고 있음. 그렇기 때문에 이미 한화생명, SK텔레콤등은 ‘4차 산업혁명 기술활용’을 광고하며 자신의 개인건강정보 결합사례를 보도까지 하는 상황임(그림1-7).
- 즉 현재의 비식별화 가이드라인은 법적 근거가 전혀 없음. 때문에 잘못된 비식별화 및 재식별화에 대한 법적 처벌의 근거도 없음. 또한 비식별화 평가를 해당기관에 위임하여 사실상 자체평가 후 빅데이터 이용이라는 방임형임. 이는 전세계적으로도 유례가 없음.
- 비식별화 정보의 결합도 해당부분 전문기관으로 공공기관이 아님(호주 PHRN등 해외는 공공기관에서 결합서비스를 제공), 또한 최근 밝혀졌듯이 결합서비스는 통신회사 등과 결합되어 재식별화의 위험성이 더욱 증가함.

이중 업종과의 개인정보 비식별 결합

다양한 업종과의 개인정보 비식별 결합을 통해 이중 데이터를 활용한 내부 프로세스 개선 가능성을 확인하고, 보험데이터의 활용가치를 검증함으로써 수익모델화 검토

대상	결합 목적	비고
한화손해보험	- 금융계열사 데이터 융합활용 시너지 측면에서 중핵가입고객의 규모 및 고객성향, 마케팅 인사이트 도출 - 결합고객의 가입채널, 가입여력, 상품 등 각 사별 변수의 영향도 분석	완료
SK텔레콤	- Alternative 신용평가모델 개발을 위한 보험/통신데이터의 신용평가 영향도 분석을 통한 개발 가능성 실증 - 고객 및 보험데이터의 유용성 및 데이터 유동 상용화의 가능성 확인	완료
신한카드	전략적 제휴 추진과제의 일환으로 보험데이터와 카드소비 데이터를 활용한 신용평가 세분화 가능성 검증	추진 중
우리은행	우리은행 자본투자 관련 시너지협업체 추진과제로 은행거래데이터와 보험거래데이터를 활용한 신용평가 검증 협의	협의 중
유통/Payment	TIMON, PAYCO 등의 거래데이터와 결합하여 소비성향과 보험/신용도 등의 연관성 검증 및 활용방안 도출	추진 중

→ 국내 최초로 2번의 개인정보 비식별 결합을 완료했으며, 한국인터넷진흥원 및 빅데이터포럼에서 다양한 빅데이터 관계사를 대상으로 사례발표 실시

Copyright © Hanuha Life. all rights reserved.



최종 결합 데이터 목록

보험사 고객정보, 거래정보, 신용정보 + 통신사 고객정보, 거래정보, 신용정보

- 총결합 데이터(42개 항목) : 한화생명(21개) + B통신사(21개)

한화생명	직업	신용대출건수	최초대출날짜	최초연체날짜	총신용대출금액	총상환금액	신용대출연체율
	최근1년 신용대출연체율	30일내 신용대출연체율	최초신용등급	최근신용등급	보험료연체율	최근1년 보험료연체율	실호해지건수
	가입일보험료	월납입보험료	직업기반 추정소득금액	가구단위 추정소득금액	평균약관대출율	약관대출금액	자동차제 실패율수
	나이	성별	사용개월수	멤버십등급	월평균통화시간	월평균통화빈도	ARPU
B통신사	결합상품가입여부	단말가솔고가	이용경치기간	당월 통신료연체금액	최근1년 최대 통신료연체금액	납부방법	최선상태
	남은단말기 할부원금	가입희선수	태블릿PC 보유여부	스마트워치 보유여부	멤버십 당월사용금액	멤버십 당년사용금액	통신료 미납횟수

준식별자 민감정보 **신용정보**

Copyright © Hanuha Life. all rights reserved.



<그림1-7> 4차산업관련 기술활용. 2017년 6월. 한화생명

7. 개인건강정보와 재식별화의 위험성

- 한국의 비합리적인 무차별적 비식별화 정보공유는 해외보다 훨씬 더 큰 치명적 개인 의료/건강 정보의 유출로 인한 신체적, 정신적, 물질적 피해가 예상됨. 이는 개인의 정신적 신체적 피해를 뿐만 아니라, 빅데이터로 다른 여러 가지 정보와 결합될 경우, 구조화된 차별 가능성까지 커짐.
- 물질적 피해도 예상. 미국의 경우 병원 의료정보 해킹 범죄가 전체 범죄의 2위가 되는 상황에 있어 FBI가 이에 대한 단속에 고심하고 있다는 연구논문도 발표되고 있음. 병원 의료정보를 해킹하고 이를 외부 유출하겠다는 협박, 돈을 요구하는 경우 등으로 최근 매우 다양한 범죄로 활용되고 있는 형국임.
- 빈번한 개인정보 유출로 인해 개인 의료/건강정보의 비밀 보장에 대한 전반적 신뢰도를 저하시켜 국내 공적 의료 시스템 전반에 대한 신뢰 하락. 이는 국가 전체 측면에서 경제적, 사회적으로 큰 손실 발생. 건강보험에 대한 불신, 보건의료제도에 대한 불신, 환자-의료인간 불신 등의 사회적 불필요한 논쟁을 유발, 보건복지 정책에 악영향을 미치게 될 것임.
- 또한 ‘거품 경제’ 활성화로 국민에게 경제적 부담 전가됨. 효과나 효용이 입증되지 않은 원격의료, 정밀의료(맞춤의료서비스), 건강관리서비스 등 각종 산업의 거품 경제 증가함.
- 부수적으로 불필요한 정보 보안 산업의 활성화도 거품경제의 한 측면으로 볼 수 있음. 정보 보안 수준이 아무리 높아지더라도 해킹의 위험으로부터 완전히 자유로울 수 없음, 결국 정보 보안 인플레이션이 작용하여 국가적 차원에서 정보 보안 수준을 기술적으로 높이려는 노력이 증가함에 따라 경제적 부담 증가할 수밖에 없음.
- 특히 최근 연구들에 따르면 보안기술이 해커기술을 따라갈 수 없다는 것이 보안 전문가들의 일치된 견해임. 해커기술은 기하급수적으로 발전하며, 이는 보안기술과 상호발전하기 때문에 서로가 서로를 강화하는 측면도 강함. 즉 정보화의 이면에는 과잉정보화로 인한 과잉보안 인플레이션을 형성. 사회적 부적절한 비용낭비를 가

저음. 결국 정보화가 필요 없는 것은 정보화하지 않아야 하고, 한곳에 집적하지 않는다는 것이 전문가들의 일치된 견해임.

- 그러므로 충분한 의사소통과 공론화를 거쳐 신뢰를 바탕으로 한 거버넌스 체계로 정책 추진이 이루어지지 않는다면, 사회적 논란과 갈등만 심화시킨 채 언제든지 좌초될 수 있는 성격의 정책임을 영국의 ‘Care.data’(케어데이터) 사업의 사례가 극명하게 보여주고 있음. 영국 NHS의 케어데이터는 최초는 장미빛 환상에서 출발하였으나, 개인건강정보의 유출, 재식별화 가능성으로 현재 폐쇄된 상태임(2016년 7월). 영국의 빅데이터사업이 국가주도의료체제의 장점으로 가장 유리할 것으로 사료되었던 것에 비추어 매우 놀라운 현상임.
- 당시 CQO(Care Quality Commission)는 보고서를 통해 투명성, 개인정보 보호 및 기밀성 보장, 추후에 동의를 철회할 수 있는 옵트 아웃(Opt-out ; 개인정보 공유금지) 권리 부족, 데이터 소유권, 상업화 등을 지목했음. 즉 투명한 빅데이터 관리가 아니었고, 개인정보축적에 대한 충분한 설명이 없었다는 뜻임. 보건의료 빅데이터에서 논란은 기본적으로 목적 외 사용의 위험성임.
- 이런 측면에서 건보공단의 정보 수집 권한은 국민건강보험을 관리하기 위한 목적으로만 부여된 것이라는 점에서 앞으로 논란도 커질 것임.

8. 대안

- 우선 가장 시급한 문제는 박근혜 정부의 비식별화 가이드라인을 철회하는 것임. 최소한 미국수준의 법률적 장치가 마련되고, 비식별화를 위한 제3기구 등이 논의되어야 함. 아니라면 유럽수준의 강력한 개인정보보안법률까지 요구됨. 개인정보에 있어서 가장 방임형인 미국, 호주에서도 상상하지 못할 규제완화와 방임형 제도는 빅데이터 산업자체에도 결코 긍정적이지 않을 것임.
- 이를 위해 법률제정이 요구됨. 이 법률은 빅데이터 산업계의 이익이 아니라, 개인 건강정보에 근거한 차별과, 개인의 건강정보 프라이버시를 보호하기 위한 제도적 장치를 만드는 것이 우선적 목적이 되어야 함. 이를 위해서 다음의 조건들이 해결되어야 함. 1) 정보소유자의 권리 보호 2) 적절한 보안조치 및 악의적 이용에 대한 대책 - 예방대책, 사후대책 3) 정보연계 보호를 위한 거버넌스 문제 4) 기관 및 수용단체 사이의 연계시 정보보호의무 문제.
- 다음으로 개인건강정보를 중심으로 만들어지는 생산물, 결과들은 각종 제약, 치료, 생활수준 및 인구집단에 대한 이익을 통해 철저하게 공적으로 관리되고 이용되어야 한다는 점임. 이는 개개인이 자신의 개인정보를 민간에게 제공했다고 하더라도, 사회적 차원에서만 이것이 데이터가 되어 확률적, 역학적 접근을 가능하게 해 주었다는 점에서 그러함. 즉 공익적으로 데이터는 관리되고 공익적으로만 활용되어야 한다는 논리적 결론이 타당함, 이를 위해서 공익적 임상시험 프로그램과 이를 통해 얻은 이익이 공공에서 확대되도록 고려되어야 함.
- 그리고 장기적으로는 생명현상을 이해하는 새로운 관점이 모색되어야 하며, 이는 개인건강정보의 집적화 및 맞춤형 서비스가 아니라 포괄적인 정치적, 사회적, 경제적 접근을 동반한 프로그램이 되어야 함. '건강결정요인의 이해'가 있어야 빅데이터의 발전적 활용이 가능할 것임. 그리고 그러한 이용은 단순한 민간기업의 이익만을 위한 것이 되어서는 안됨.

[보론] 법률상 개인정보 등의 현재 문제와 상충되는 조항들

개인정보 보호법 제18조 - 원칙적으로 개인정보는 수집 목적 범위를 초과하여 제3자에게 제공할 수 없도록 되어 있고 단지 아래와 같은 경우만 예외로 인정하고 있음.

1. 정보주체로부터 별도의 동의를 받은 경우
2. 다른 법률에 특별한 규정이 있는 경우
3. 정보주체 또는 그 법정대리인이 의사표시를 할 수 없는 상태에 있거나 주소불명 등으로 사전 동의를 받을 수 없는 경우로서 명백히 정보주체 또는 제3자의 급박한 생명, 신체, 재산의 이익을 위하여 필요하다고 인정되는 경우
4. 통계작성 및 학술연구 등의 목적을 위하여 필요한 경우로서 특정 개인을 알아볼 수 없는 형태로 개인정보를 제공하는 경우
5. 개인정보를 목적 외의 용도로 이용하거나 이를 제3자에게 제공하지 아니하면 다른 법률에서 정하는 소관 업무를 수행할 수 없는 경우로서 보호위원회의 심의·의결을 거친 경우
6. 조약, 그 밖의 국제협정의 이행을 위하여 외국정부 또는 국제기구에 제공하기 위하여 필요한 경우
7. 범죄의 수사와 공소의 제기 및 유지를 위하여 필요한 경우
8. 법원의 재판업무 수행을 위하여 필요한 경우
9. 형(刑) 및 감호, 보호처분의 집행을 위하여 필요한 경우

개인정보 보호법 제23조 - “사상·신념, 노동조합·정당의 가입·탈퇴, 정치적 견해, 건강, 성생활 등에 관한 정보, 그 밖에 정보주체의 사생활을 현저히 침해할 우려가 있는 개인정보로서 대통령령으로 정하는 정보”는 “민감정보”로 규정하고, 이러한 민감정보는 별도의 동의를 받거나 법령에서 민감정보의 처리를 요구하거나 허용하는 경우 외에는 처리를 제한하고 있음.

2017년 국감에서 정의당 추혜선 의원실에 따르면

- 1) ‘빅데이터 활용’을 활성화한다는 명분으로 박근혜 정부 당시 추진됐던 ‘개인정보 비식별 조치 가이드라인(이하 비식별 가이드라인)’에 따른 정보집합물 결합 서비스를 통해 3억 4천여만 건의 개인정보가 공공기관을 통해 기업들에 제공’.
- 2) 한국인터넷진흥원, 정보통신산업진흥원, 금융보안원, 한국신용정보원 등 비식별 전문기관 4곳으로부터 입수한 자료에 따르면, 비식별 가이드라인 도입(2016.6) 이후 현재까지 이들 기관이 총 26건의 정보집합물 결합서비스를 통하여 3억 4천여만 건의 개

인정보 결합물을 기업 등에 제공한 사실이 확인됨.

3) 삼성생명과 삼성카드는 2017년 2월 양사에 동시 가입한 240만여 고객의 ‘가입건수, 보험료, 가입기간, 가입상품 및 카드이용실적정보’등의 개인정보를 13회에 걸쳐, 보험개발원과 현대자동차는 1억 5천만 건의 고객정보를 두 차례에 걸쳐 ‘비식별 가이드라인’에 따른 ‘정보집합물 결합지원서비스’를 통해 결합 제공받았음.

26개 인종 6000명의 유전체 정보 분석... ‘바이오 구글’ 서막, 맞춤의료 시대 이끈 신테카바이오 - 한국일보 2017년 1월 5일

[Bio & Tech] AI로 빅데이터 분석...맞춤형 항암제 효능도 미리 알아볼 수 있어요.

신테카바이오 - 매일경제 2017년 9월 6일

4) 박근혜 정부가 ‘국가중점개방 데이터 공개’라는 명목으로 개인 및 기업에게 제공하고 있는 국민 개인의 진료내역, 약품처방, 건강검진 내역은 아래 어느 사항에 해당하여 제3자에게 제공하고 있는 것인지 불분명함.

개인정보 비식별화의 문제점

이은우 / 정보인권연구소 이사, 법무법인지향 변호사

1. 건강정보의 보호

1) 건강정보는 민감정보

- 개인의 병력이나 질병, 현재의 건강상태 등에 관한 정보는 개인정보 중 가장 민감한 정보 중 하나(제23조).
- 공개될 경우 개인에 대한 사회적 낙인이 될 수 있음.
- 고용, 보험은 물론 사회생활에서 차별의 원인이 될 수도 있음.
- 개인의 민감한 사생활 침해가 될 수도 있음.
- 의료법, 약사법 등에서도 진료기록과 처방에 대한 기록을 엄격하게 보호.

2) 건강정보의 활용

- 환자 개인에게는 질 높은 치료를 제공하기 위해서 필요.

- 환자들이 소통, 투명, 경험, 응답 등에 필수적.
- 공공의 목적으로는 새로운 치료법과 연구개발을 위해 필요.
- 보건 관련 재정의 제한으로 인한 보건정책 추진에 필수적.

3) 건강정보의 활용은 이와 같은 공공 목적을 위해서 필요최소한의 범위로, 엄격한 법적 보호조치를 취한 상태로 이루어져야 함

2. 건강정보에 대한 엄격한 ‘개인정보보호법’ 과 ‘의료법’ 의 보호

1) 개인정보보호법은 개인의 건강에 관한 정보를 민감정보로 규율(제23조 제1항)

- 민감정보의 처리는 원칙적으로 금지되고, 정보주체에게 고지할 사항을 알리고 다른 개인정보의 처리에 대한 동의와 별도로 동의를 받은 경우이거나, 법령에서 민감정보의 처리를 요구하거나 허용하는 경우에만 처리가 허용된다(제23조 제1항 제1호, 제2호).

2) 개인정보보호 원칙 준수

- 민감정보의 수집, 처리시에는 당사자의 동의를 받았거나 법률의 규정이 있다고 하더라도 최소수집의 원칙, 익명 처리의 원칙 등 개인정보 보호 원칙과, 정보를 제공받을 권리, 접근권, 열람 및 정정, 삭제, 처리중지 요구권, 철회권 등의 정보주체의 권리는 보장되어야 함. 침해시에는 권리구제를 받을 수 있음.

3) 의료인이나 의료기관은 엄격한 비밀준수 의무를 부담하고 있으나, 건강보험공단과 심평원은?

- 의료법(제19조, 제69조), 약사법(제87조), 정신보건법(제71조) 등의 비밀보호 의무.
- 형법상 업무상비밀누설죄(제371조 제1항)로 직무상 알게 된 비밀의 누설이나 공표를 엄격하게 규율.

4) 의료인이나 의료기관은 환자의 동의 없이 제3자 제공을 할 수 있는 경우를 한정적으로 열거하고 있는데, 건강보험공단과 심평원은?

- 의료법과 약사법은 환자정보를 당사자의 동의가 없어도 제3자에게 제공할 수 있는 예외를 한정적으로 열거(의료법 제21조 제2항, 제3항, 약사법 제30조 제3항).

<표2-1>의료인이나 의료기관이 환자의 동의 없이 제3자에게 제공할 수 있는 경우

요건	제3자
급여비용 심사·지급·대상여부 확인·사후관리 및 요양급여의 적정성 평가·가감지급 등을 위하여	국민건강보험공단 또는 건강보험심사평가원
의료급여 수급권자 확인, 급여비용의 심사·지급, 사후관리 등 의료급여 업무를 위하여	보장기관(시·군·구), 국민건강보험공단, 건강보험심사평가원
보험급여를 받는 근로자를 진료한 산재보험 의료기관(의사를 포함한다)에 대하여 그 근로자의 진료에 관한 보고 또는 서류 등 제출을 요구하거나 조사하는 경우	근로복지공단
의료기관으로부터 자동차보험진료수가를 청구 받은 보험회사 등이 그 의료기관에 대하여 관계 진료기록의 열람을 청구한 경우	의료기관으로부터 자동차보험진료수가를 청구 받은 보험회사 등
병역판정검사와 관련하여 질병 또는 심신장애의 확인을 위하여 필요하다고 인정하여 의료기관의 장에게 병역판정검사대상자의 진료기록·치료 관련 기록의 제출을 요구한 경우	지방병무청장
부양가족연금, 장애연금 및 유족연금 급여의 지급심사와 관련하여 가입자 또는 가입자였던 사람을 진료한 의료기관에 해당 진료에 관한 사항의 열람 또는 사본 교부를 요청하는 경우	국민연금공단
공무상요양비, 재해보조금, 장애급여 및 유족급여의 지급심사와 관련하여 공무원 또는 공무원이었던 자를 진료한 의료기관에 해당 진료에 관한 사항의 열람 또는 사본 교부를 요청하는 경우	공무원연금공단
장애 정도에 관한 심사와 관련하여 장애인 등록을 신청한 사람 및 장애인으로 등록한 사람을 진료한 의료기관에 해당 진료에 관한 사항의 열람 또는 사본 교부를 요청하는 경우	공공기관의 장
감염병의 역학조사 및 예방접종에 관한 역학조사를 위하여 필요하다고 인정하여 의료기관의 장에게 감염병환자등의 진료기록 및 예방접종을 받은 사람의 예방접종 후 이상반응에 관한 진료기록의 제출을 요청하는 경우	보건복지부장관, 질병관리본부장, 시·도지사 또는 시장·군수·구청장

3. 국민건강보험공단과 심평원의 건강정보 보유 및 이용실태는 어떤가?

1) 의료관련 사회보장제도의 운영과 관련하여 건강정보 보유, 활용하는 건보공단과 심평원

<표2-2> 우리나라 사회보장제도

분류	해당 사회보장제도	
사회보험	• 국민연금 • 노인장기요양보험 • 산업재해보상보험	• 건강보험 • 고용보험
공공부조	• 생계급여 • 주거급여	• 의료급여 • 교육급여
사회서비스	• 노인복지서비스 • 아동복지서비스 • 가정복지서비스	• 장애인복지서비스 • 영유아복지서비스

<표2-3> 의료 관련 사회보장제도

보험제도	보험자(보장기관)	관련법률
국민건강보험	국민건강보험공단	국민건강보험법
의료급여	지방자치단체장(특별시장·광역시장·도지사·시장·군수·구청장)의 업무 중 수급권자의 관리, 급여비용의 심사·조정 및 지급 업무 등 의료급여에 관한 업무를 국민건강보험공단에 위탁	의료급여법
노인장기요양보험	국민건강보험공단	노인장기요양보험법
산업재해보상보험	근로복지공단	산업재해보상보험법

○ 국민건강보험공단은 건강보험과 노인장기요양보험의 보험자, 의료급여의 수탁기관, 국민건강증진기관으로서 역할을 수행하는 과정에서 다음과 같은 건강정보를 취급하고 있음.

- 가입자 및 피부양자의 자격 관리를 위해 필요한 자료
- 보험료 부과, 징수를 위해 필요한 자료
- 보험급여의 관리를 위해 필요한 자료
- 보험급여 비용의 지급을 위해 필요한 자료
- 건강 증진과 예방사업을 위해 필요한 자료를 광범위하게 수집

- 심평원은 다음과 같이 심사평가 업무를 담당하고 있음.
- 요양급여비용을 심사하고 요양급여의 적정성을 평가하기 위해 설립된 법인(국민건강보험법 제62조).
 - 심사기준 및 평가기준의 개발, 요양급여비용 심사, 요양급여 적정성 평가 업무와 관련된 조사연구 및 국제협력, 의료급여법, 노인장기요양보험법에 의해 위탁받은 급여비용 심사와 의료의 적정성에 대한 평가 업무, 심사청구와 관련된 소프트웨어의 개발·공급·검사 등 전산 관리, 요양급여의 적정성 평가 결과의 공개 업무(제63조, 시행령 제28조).
 - 국민건강보험법 제14조, 제47조, 제48조 및 제63조에 의하여 의료기관과 약국은 급여비용 심사·지급·대상여부 확인·사후관리 및 요양급여의 적정성 평가·가감지급 등을 위하여 건강정보를 국민건강보험공단 또는 건강보험심사평가원으로 제공하여야 함. 의료급여법도 요양기관이 요양급여비용을 청구하기 위해서 공단과 심사평가원에 요양급여의 내역을 제공하도록 하고 있고, 노인장기요양보험법도 공단과 심사평가원에 내역을 제공하도록 되어 있음.

4. 건보공단과 심평원, 질병관리본부 등이 수집하는 정보는 적법하고, 신뢰할 수 있는가?

<표2-4> 보건복지분야 국가승인통계 현황

통계 종류	분야	통계 명칭
조사(42종)	보건(23)	건강보험환자 진료비 실태조사, 국민건강영양조사, 국민보건의료실태조사, 근로환경조사, 병원경영실태조사, 아동구강건강 실태조사, 의료기관별 급여 적정성 평가현황, 의료기기 제조/유통 조사, 의약품/의료기기연구개발실태조사, 작업환경실태조사, 전국민장내기생충감염실태조사, 전국출산력 및 가족보건복지실태조사, 정신질환실태조사, 지역사회건강조사, 청소년건강행태 온라인조사, 퇴원손상심층조사, 한국의료패널조사, 한국인인체치수조사, 한방의료이용 및 한약소비실태조사, 한의약산업실태조사, 화장품제조유통조사, 환자조사
	복지(19)	가정폭력실태조사, 고령화연구패널조사, 국민노후보장패널조사, 기업 및 공공기관의 가족친화수준조사, 남해군 노인실태조사, 노인실태조사, 노후준비실태조사, 보육실태조사, 복지욕구조사, 사회서비스수요/공급실태조사, 생명보험성향조사, 서울특별시 복지실태조사, 성폭력실태조사, 아동종합실태조사, 장애인생활체육실태조사, 장애인실태조사, 장애인편의시설설치현황조사, 한국복지패널조사, 한부모가족실태조사
보고(34종)	보건(21)	HIV/AIDS 신고현황, 건강검진통계, 건강보험 주요 수술통계, 건강보험통계, 결핵현황, 공중위생관계업소 실태보고, 근로자건강진단실시상황보고, 급성심장정지조사, 노인장기요양보험통계, 법정감염병발생보고, 보건소 및 보건지소운영현황, 수입식품현황, 식품 및 식품첨가물 생산실적, 식품수거검사실적, 암등록통계, 완제의약품 유통정보통계, 응급의료현황통계, 의료기기생산실적, 전국예방접종률조사, 지역별의료이용통계, 학생건강검사통계보고
	복지(13)	가정위탁 국내 입양소년소녀 가장현황, 국민기초생활보장 수급자현황, 국민연금통계, 노인복지시설현황, 노인학대현황, 보훈보상금지급현황, 산업재해현황, 산재보험통계, 아동복지시설 보호아동 및 종사자현황보고, 어린이집 및 이용자통계, 요보호 아동현황보고, 장애인현황, 학대피해 아동보호현황
가공(5종)	보건(4)	국민의료비추계 및 국민보건계정, 사망원인통계, 어린이 식생활안전지수, 의약품소비량 및 판매액 통계
	복지(1)	한국의 사회복지지출

<표2-5> 주요 빅데이터 현황

구분	보유기관	내용	공개여부
건강보험표본 코호트DB	국민건강보험공단	• 자격DB: 건강보험가입자 및 의료수급권자의 성, 연령, 대, 지역, 사회경제적 변수, 장애, 사망관련 등 • 진료DB: 요양급여 청구자료로서 진료, 상병, 처방 관련 변수 • 건강검진DB: 검진 주요결과 및 문진에 의한 생활습관 및 행태관련 자료 • 요양기관DB: 요양기관 종별, 설립구분, 지역, 시설, 장비, 인력관련 자료	제한적 공개
환자데이터셋	건강보험심사평가원	• 건강보험 청구자료를 기초로 진료개시일 기준 1년 간 진료 받은 환자대상의 표본 데이터	제한적 공개
한국인체자원	질병관리본부	• 공여자로부터 기증받은 인체유래물(DNA, 조직, 혈액, 뇨 등)과 임상(진단명, 수술명, 병리조직검사결과, 혈액검사 등). 역학(성별, 생년월일, 음주력, 흡연력 등) 및 유전(SNP, CNV, Exome 등)정보	제한적 공개
지역보건의료정보	사회보장정보원	• 전국 보건기관(보건의료원, 보건소/지소, 보건진료소)의 보건사업 및 행정업무, 전자의무기록 및 진료관련(진료내역 및 검진결과 등) 정보	미공개
지역사회건강조사	질병관리본부	• 지역 보건의료계획수립 및 보건사업 평가 활용 지표로서, 건강행태, 건강검진 및 예방접종, 질병이환, 의료이용, 사고 및 중독, 활동제한 및 삶의 질, 보건기관 이용, 사회 물리적 환경, 심정지, 교육 및 경제활동 등	공개
국민건강조사	질병관리본부	• 국민의 건강 및 영양 상태에 관한 현황 및 추이 파악 신체계측, 비만, 고혈압 등 검진조사, 흡연, 음주, 비만 및 체중조절, 신체활동 등 건강설문조사, 식품 및 영양소 섭취현황, 식생활행태, 식이보충제 등 영양조사	공개
한국의료패널	한국보건사회연구원	• 개인의 건강수준, 의료이용 및 의료비 지출 요인, 건강행태, 의료욕구, 보건의료서비스 수요행태 변화분석 사회경제적 특성, 의약품 구매, 경제활동, 건강수준, 의약품 복용행태, 민간의료보험, 건강기능식품, 건강행태 등	공개
사회보장정보	사회보장정보원	• 각 부처 및 정보보유기관에서 제공하고 있는 복지사업정보 및 지원대상자의 자격정보, 수급이력정보를 통합관리 • 복지대상자 선정·사후관리를 위해 45개 기관 552종의 소득·재산자료 및 서비스 이력정보 연계	미공개
복지콜센터 상담 데이터	복지콜129센터 등	• 상담이력	미공개
사회서비스 전자바우처	사회보장정보원	• 정부와 지자체가 사회로부터 도움을 필요로 하는 사람에게 돌봄, 일상생활 지원, 사회적응지원, 문화체험 등의 서비스를 제공하는 것을 전산처리	미공개
보육통합정보	사회보장정보원	• 보육바우처 운영 및 행정지원 정보	미공개
한국복지패널	한국보건사회연구원	• 빈곤층, 근로빈곤층, 차상위층 등의 규모 및 생활실태 변화를 동태적으로 파악 • 인구집단별 생활실태 및 경제활동, 건강·의료, 주거, 소득, 복지서비스 이용, 복지욕구	공개

1) 주요 빅데이터의 법적 근거가 명확한가?

<표2-6> 심평원의 요양급여 청구명세서 정보

항목명	세부내용
명세서조인키	명세서 레벨의 팩트간 조인 Key(각 자료간 연결키)
보험자코드	4: 건강보험, 5:의료급여, 7: 국비보훈(상이처, 무자격자)
수진자식별대체키	주민등록번호 대체번호
성별구분	9:기타, 1:남, 2:여
수진자연령	0.2: 28일미만, 0.9: 28일이상1세미만, 999: 연령산출 불가
수진자통계연령	매년 말일 기준의 수진자 연령
수진자구분코드	0:일반, 1:신생아, 2:행려
요양기관식별대체키	요양기관기호 대체번호
요양종별코드	01 상급종합병원, 11 종합병원, 21 병원, 28 요양병원, 31 의원, 41 치과병원, 51 치과의원, ... 81:약국
시도코드	11 서울, 21 부산, 22 인천, 23 대구, 24 광주, 25 대전, 26 울산, 31 경기, 32 강원, 39 제주
지역형태코드	01:특별시, 02:광역시, 03:구가있는시, 04:시, 05:군, 06:특별시의군, 07:광역시의군, 08:구가있는시군, 09:시의군
서식구분코드	021: 의과입원 031: 의과외래 041: 치과입원 051: 치과외래 ... 121: 한방입원 131: 한방외래
주상병코드	주된 상병분류기호(약국 증상분류기호, 한방상병코드, 의과:표준질병사인분류코드), (배제진단 제외)
부상병코드	주된 상병분류기호외의 추가 상병분류기호 DW에서 생성 된 ABC구분추가(배제진단 제외)
진료과목코드	요양기관에서 청구시 기재한 진료과목코드, 의원급 표시과목코드
표시과목코드	의원급의 경우 요양기관현황의 표시과목코드, 병원급이상은 청구된 진료과목코드 기준
내원일표시코드	요양기관에 실제 내원한 일자. 31 occurs.(외래서식의 경우 내원일자 위치에 9 표시)
당월요양개시일자	진료를 받기 시작한 일자, 조제투여일자, 최초내방일자(청구명세서 단위)
요양종료일자	수진자가 진료 받기를 종료한 일자

- 청구 심사 후에도 이를 보유하는 것은 적법한가?
- 보유의 법적 근거는 무엇인가?
- 식별자를 그대로 두는 것은 적법한가?
- 다른 정보(건강검진 정보, 자격정보, 징수정보 등)와 연계하는 것은 적법한가?
- 폐기하지 않고 보유하는 기간은 적절한가?

<표2-7> 기타 요양급여 정보 DB 등

개인정보파일명	개인정보 내역	파일 운영목적
의료급여내역 10년 661,740,265명	이름:필수, 주민번호:필수, 기타, 기타 (지급대상, 직역, 인정등급, 장기요양관리번호, 증번호, 서비스 종별, 청구관련 자료, 심사결정요양급여비용총액, 심사결정본인부담금, 진료일수) 지급년월, 지급일련번호, 상세지급일련번호, 수급자성명, 서비스종류, 서비스시작일, 서비스종료일, 서비스횟수, 수가코드, 수가명	의료급여, 보험급여, 보험급여비용의 지급
현금급여비 지급내역 (본인부담금환급금, 본인부담액보상금, 현금급여 포함) 10년 2,006,024 건	이름:필수, 집연락처:필수, 핸드폰(연락처):필수, 생년월일:필수, 주민번호:필수, 건강, 기타(계좌번호, 금융기관, 내구연한, 민원구입액, 보장구구입일자, 보장구구입처,보장구코드,본인부담액,사고경위,사망구분,사망 상병,사망일,사업장기호,사업장명,사용제품종류,상병코드,수진 자명,수진자주민번호,실소요액,실종자여부,심사결정액,예금주 관계,예금주성명,예금주주민번호,요양기관,의뢰일,자격상실여 부,장애등급,장애발생경위,장애발생유형,장애종류,접수유형,접 수일,중복장애여부,증번호,지급방법,지급상한액,지급액,처리지 사,청구사유,청구인관계,청구인성명,청구인전화번호,청구인주 민번호,청구인주소,최초지급,최초지급,출산구분,출산일,출산자 명,출산자주민번호,출산장소,투약개시일,투약기간)	현금급여비 지급내역 관리로 지급대상자에게 적정 지급여부 확인
임신·출산 진료비 지원신청내역 5년 3,256,198명	이름:필수, E-Mail:필수, 집연락처:필수, 핸드폰(연락처):필수, 주민번호:필수, 외국인등록번호:필수, 건강:필수, 기타(수진자 성명, 전화번호, 휴대폰번호, 임신확인일, 사실확인일, 분만에정일, 다태아구분, 요양기관기호, 요양기관명, 의사면허번호, 담당의사명, 신청자 성명, 수진자와의 관계, 신청인 전화번호, 신청일자)	임신·출산 진료비지원 대상자의 급여내역 및 카드발급 관리
요양급여내역 10년 11,162,632,450명	이름, 주민번호, 외국인등록번호, 건강, 성생활, 유전정보, 기타, 기타(청구서내역, 명세서내역, 심사조정내역)	건강보험 지급업무의 관리·운영
요양급여비용청구명세서 5년 7,374,850,644명	이름, 주민번호, 외국인등록번호, 건강, 성생활, 유전정보, 기타, 기타(청구서내역, 명세서내역, 진료내역, 상병내역, 처방전교부내역, 처방전상세내역, 심사조정내역)	보험급여 사전·사후관리 활용

- 청구 심사 후에도 이를 보유하는 것은 적법한가?
- 보유의 법적 근거는 무엇인가?
- 식별자를 그대로 두는 것은 적법한가?
- 다른 정보(건강검진 정보, 자격정보, 징수정보 등)와 연계하는 것은 적법한가?
- 폐기하지 않고 보유하는 기간은 적절한가?

<표2-8> 건강검진 등과 관련 수집, 이용하는 건강정보

개인정보파일명	개인정보 내역	파일 운영목적
건강검진 대상자 및 검진결과 내역 영구 722,523,512명	이름:필수, 집주소:필수, E-Mail:필수, 집연락처:필수, 핸드폰(연락처):필수, 주민번호:필수, 외국인등록번호:필수, 건강, 기타(대상자·수검자·검진의사 성명 및 주민번호, 일반·암검진·영유아 검진결과 및 종합판정, 문진내역, 소견 및 조치, 검진기관명, 검진의사 면허번호, 검진(판정)일, 개인정보활용동의서)	가입자 및 피부양자 건강검진 실시 및 평생 건강관리 체계 수행
검진사후관리 대상자 내역 10년 77,134,811명	이름:필수, 집주소:필수, 직장주소:필수, E-Mail:필수, 집연락처:필수, 직장연락처:필수, 핸드폰(연락처):필수, 생년월일:필수, 주민번호:필수, 건강:필수, 기타, 기타(가입자 및 세대주 정보(보험료, 성명, 주민번호, 관계코드, 취득일자, 동거가족구분, 성별, 연령, 건강보험증번호, 가족수), 결혼상태, 교육정도, 직업, 사업장기호, 차상위 여부, 차상위구분, 차상위종별, 1차일반건강검진및생애전환기건강진단결과, 2차일반건강검진및생애전환기건강진단결과, 암검진결과, 대사증후군 상담관리, 유질환(집중관리)군 상담관리, HCA발급관리, HRA발급관리)	고혈압, 당뇨병, 대사증후군 등 만성질환자 및 질환 전단계 대상자의 건강관리, 건강증진, 생활습관개선, 보건교육 등 예방사업 수행
합리적의료이용지원 및 적정투약관리 대상자 관리내역 10년 8,092,626명	이름:필수, 집주소:필수, 직장주소:필수, E-Mail:필수, 집연락처:필수, 직장연락처:필수, 생년월일:필수, 주민번호, 건강:필수, 기타 (가루약종, 가입자 구분, 가입자 주민번호SEQ, 가입자와의 관계, 거주지지사, 계단 오르기, 공단부담금, 내원일수, 단골병의원명, 단골약 국명, 대상자 동거가족, 대상자 성별, 대상자 연령, 목록하기, 보건복지부고시번호, 복용약물 수의 감소, 복용약물 종류의 감소 방지, 생활습관기타, 소속지사, 수진자 성명, 수진자 주민번호, 수진자 주민번호SEQ, 시럽종, 식사량, 식사정도, 신장, 알, 알약종, 약국급여비, 약국이용이관 수, 약국조제일수, 약국진료비, 약국진료일수, 약국투약일수, 옷 벗고입기, 외래 내원일수, 외래 진료일수, 외래급여비, 외래이용기관 수, 외래이용일수 감소, 외래진료건수, 외래진료비, 외래진료일수, 요양기관 중복경로, 요양기호, 우편번호, 운동방법, 운동정도, 운동횟수, 월보험료, 음주량, 음주정도, 의료기관수 감소, 의료기관이용양상 기타, 일상생활수행능력 기타특이사항, 임원급여비, 임원이용기관 수, 임원일수, 임원진료건수, 임원진료비, 임원진료일수, 중복간격(일), 중복투약일수, 증번호, 지역구분, 진료일자, 처방(조제)기관 2, 처방(조제)기관1, 처방(조제)기관명1, 처방(조제)기관명2, 처방(조제)일자, 체중, 총의료기관 이용기관 수, 피면담자 관계, 혼자잠자리살피기, 휴대전화번호, 흡연량, 흡연정도, BMI지수,cc, 부상병, 비만상태, 상담의사 진료과명, 성분명1(영문), 성분명1(한글), 성분명 2(영문), 성분명2(한글), 약물 오·남용 이력, 약품명1, 약품명2, 약품코드1, 약품코드2, 주상병, 298상병)	가입자의 합리적 의료이용지원 및 적정투약관리(건 강유지 및 증진 위하여 필요한 예방사업)
만성질환자 건강지원서비스 준영구 350,462,071명	이름:필수, 집주소:필수, 직장주소:필수, E-Mail:필수, 집연락처:필수, 직장연락처:필수, 핸드폰(연락처):필수, 생년월일:필수, 주민번호:필수, 건강:필수, 생활:필수, 유전정보:필수, 기타:필수, 기타(등록자관리번호,등록자건강보험증번호,등록자관리지사, 서비스신청일자, 등록자사업장기호, 등록자보험료, 등록자면담방법, 등록자가입경로, 등록자진행상태, 등록자특이사항, 자가측정기 신청여부, 검진사후이관여부, 의원급만성질환관리제참여여부, 지역자원보건소여부,	만성질환자의 건강 유지 및 증진을 위한 건강지원서비스 서비스 제공

	제3자정보제공동의여부, 서비스제공면담결과구분, 면담일시, 기본상담일시, 면담상세내역, 교육강사자격구분, 교육강사계약기간, 교육신청내역, 교육리더여부, 측정기대어 및 반납일자, 서비스진행상태, 발송SMS내용, SMS발송전화번호, 안내문발송일자, 동거가족유무, 직업유무, 직업코드, 요구사정문항별 답변내용, 문제목록명, 면담내용, 피상담자반응구분, 개인별수행활동내용)	
--	---	--

○ 이 정보들을 보유하는 법적 근거는 무엇인가?

○ 식별자를 그대로 두는 것은 적법한가?

○ 다른 정보(건강검진 정보, 자격정보, 징수정보 등)와 연계하는 것은 적법한가?

○ 폐기하지 않고 보유하는 기간은 적절한가?

<표2-9> 국민건강보험 DB(DW 건강보험통계 DB)

항목
주민번호, 건강, 기타 ○ 건강보험 적용인구, 세대정보 : 주민등록번호, 장애등급, 장애유형, 장애등록일자, 체류자격, 국가, 주소, 실거주주소, 건강보험증번호, 사업장, 행정전산망세대주민등록번호, 국가유공자, 국가유공자자료구분, 장애인세대여부, 단독세대여부, 외국인세대구분, 만성희귀질환세대여부, 건강보험적용 세대수, 개인성명, 출입국일자, 출국사유, 출입국구분, 전역여부, 군기관유형, 군복무유형, 군전환복무여부, 군입대년월, 군전역년월, 국가유공자주민등록번호, 국가유공자명, 국가유공자등록일자, 국가유공자제외일자, 상이등급세부사항, 자격취득일자, 자격상실일자, 자격취득구분, 상실구분, 세대주구분, 보험급여정지사유, 시도, 시군구, 행정동, 법정동, 리○ 사업장 및 지역보험료 부과정보, 징수, 체납 정보 : 사업장, 사업자등록번호, 개업일자, 대표자주민등록번호, 사업자소득구분, 종합소득종류, 사업자소득금액, 건강보험증번호, 주소, 총소득금액, 자동차최초등록일자, 자동차연식구분, 변경일자, 지역보험료자동차번호, 차종, 자동차용도, 자동차배기량, 자동차적재량, 승차정원수, 자동차연식구분, 자동차세액, 총재산금액, 보험료수납일자, 총납부금액, 징수당월보험료납부금액, 과년도보험료납부금액, 환급지급보험료, 환급지급가산금, 보험료감감, 보험료면제구분, 직장보수월액, 직장산정보험료, 직장경감보험료, 직장면제보험료, 직장조정특례감면보험료, 직장가입자분할고지보험료, 직장수시정산추가보험료, 직장수시정산반환보험료, 직장종직정산보험료, 직장중간정산보험료, 직장연말정산보험료, 직장본인부담금환급금, 직장가입자부담보험료, 직장사업장부담보험료, 직장총고지보험료, 직장국가부담보험료, 건강보험장기요양구분, 상계총당금액, 법인상계총당금액, 상계총당보험료, 직장납부보험료, 직장납부가산금, 직장보험료과오납금액, 직장가산금과오납금액, 국가부담납부보험료, 지역정기산정보험료, 지역산정보험료, 지역경감보험료, 지역정지보험료, 지역면제보험료, 지역도서벽지경감금액, 지역농어촌경감금액, 지역농어업인경감금액, 지역세대경감금액, 지역경감반영보험료, 지역정기부과보험료, 지역정산반영보험료, 지역정산부과보험료, 지역고지보험료, 지역최종부과보험료, 지역총체납보험료, 결손처분보험료, 압류상태, 압류유형구분, 압류일자, 압류취소사유, 압류취소일자○ 일반건강검진, 암검진, 생애전환기 검진 : 주민등록번호, 건강검진유형코드, 외국인여부ID, 체류자격코드, 국가코드, 건강검진일자, 검진비지급일자, 장애유형코드, 장애등급코드, 의사면허번호, 의사주민등록번호, 일반 건강검진 결과코드, 일반건강검진 문진결과코드, 생애전환기 검진결과코드, 생애전환기 문진결과코드, 생애전환기 암검진 결과코드, 생애전환기 건강검진암검사문진 결과코드, 영유아건강검진 대상자, 과거병력 문진내역코드, 영유아검진구강검사 결과코드, 건강검진구강검사 결과 코드, 건강검진대상자코드, 암검진대상자코드, 의료급여암검사대상자코드, 생애전환기건강검진대상자코드○ 노인장기요양 인정신청 및 급여 : 장기요양관리번호, 개인ID, 주민등록번호, 인정여부, 1차판정등급코드, 인정등급코드, 신청일자, 대상자구분코드, 인정등급판정일자, 의료급여보장기관코드, 의료급여시설기관코드, 의료급여종별코드, 의료급여유형코드, 거주지사시도시군구코드, 장애등급코드, 장애유형코드, 장애등록일자, 체류자격코드, 국가코드, 지역상세코드, 동거여

부ID, 동거인관계코드, 신청자건강보험증번호, 가입자구분코드, 만성희귀질환세대여부ID, 산정보험료, 보수월액, 도서벽지거주여부, 자격취득일자, 자격취득사유코드, 자격상실일자, 상실사유코드, 사망여부ID, 사업장관리번호, 주거형태코드, 동거인여부, 동거인유형별코드, 자격취득구분코드, 실거주지주소코드, 노인장기요양신청일, 장기요양보호자관계코드, 인정변경신청사유코드, 노인성질병코드, 장기요양변경전등급분류코드, 유효기간시작일자, 유효기간종료일자, 대리인유형구분코드, 대리인신청관계기타코드, 의사소견서발급기관기호, 의사소견서발급구분코드, 요양기관진료과목코드, 의사소견서치매노인코드, 인정조사치매여부, 종전등급코드, 장기요양인정점수, 건강보험재산상태코드, 월소득금액, 주요질병코드, 장기요양주요질병코드, 감경적용기초코드, 감경적용여부, 2차판정급여종류, 질환 여부, 갱신이전인정등급코드, 판정결과, 장기요양인정조사표코드, 의사소견서코드, 장기요양기관 인력현황코드, 요양보호사발급대상코드, 복지용구 정보, 한도관리내역, 이용상담내역,급여계약내역서 정보, 이용지원 상담내용을 관리, 노인장기요양급여비용 명세서코드, 노인장기요양급여비용 서비스내역코드○ 요양급여비용 지급 및 중증암환자 등록, 현금급여 지급 : 건강보험증번호, 주민등록번호, 현금급여지급현황코드, 급여사후결정코드, 가입자피부양자구분, 장애등급, 장애유형, 체류자격, 외국인여부ID, 주소, 사업장, 진료결과, 요양급여구분, 진료과목, 공상산재구분, 수술구분, 정액정률구분, 진료개시일자, 요양기관기호, 심사결정DRG번호, 요양급여입내원일수, 요양급여투약일수, 심사결정총진료비, 심사결정보험자부담금, 심사결정본인부담금, 심사결정보험자부담금차감액, 심사결정본인부담액상한초과금액, 심사결정식대기본금액, 심사결정식대가산금액, 본인부담금환급금, 심사조정건수, 요양급여비용총액, 산정특례대상자 등록번호, 주상병, 부상병, 공무상요양비지급코드, 중증암등록환자코드○ 의료급여 적용, 세대, 급여비용 : 주민등록번호, 국가유공자, 국가유공자자료구분, 의료급여종별, 의료급여유형, 행정동의료보호기관, 의료급여증번호, 세대주주민등록번호, 세대주나이, 세대주성별, 장애등급, 장애유형, 체류자격, 외국인여부ID, 주소, 진료결과, 요양급여구분, 진료과목, 공상산재구분, 수술구분, 정액정률구분, 진료개시일자, 요양기관기호, 심사결정DRG번호, 요양급여입내원일수, 요양급여투약일수, 심사결정총진료비, 심사결정보험자부담금, 심사결정본인부담금, 심사결정보험자부담금차감액, 심사결정본인부담액상한초과금액, 심사결정식대기본금액, 심사결정식대가산금액, 본인부담금환급금, 심사조정건수, 요양급여비용총액, 산정특례대상자 등록번호, 주상병, 부상병)

- 보존기간 준영구

- 파일 운영 목적 : 국민건강 증진과 건강보험의 지속가능성 관련 연구목적)

○ 이 정보를 보유하는 법적 근거는 무엇인가?

○ 식별자를 그대로 두는 것은 적법한가?

○ 다른 정보(건강검진 정보, 자격정보, 징수정보 등)와 연계하는 것은 적법한가?

○ 폐기하지 않고 준영구로 보유하는 기간은 적법한가?

<표2-10> 코호트 DB

명칭	대상	추적기간	자료건수 (천 건)	내용
표본코호트DB	2006년 1년간 건강보험가입자 및 의료급여수급권자 자격을 유지한 전국민 중 2% 추출(100만 명) 전국민 모집단의 2%, 성·연령·가입자구분·보험료 분위·지역별 층화추출	2002 ~ 2015년(14년)	2,619,397	사회·경제적 현황(자격 및 보험료, 장애 및 사망), 의료이용 현황(진료 및 건강검진), 영양기관 현황
건강검진코호트DB	2002년 자격유지자 중 2002~2003년 40~79세 일반건강검진 수검자 중 약 51만 명(모집단의 10%)	2002 ~ 2015년 (14개년)	2,087,629	사회·경제적 자격 변수(장애 및 사망 포함), 의료이용(진료 및 건강검진)현황, 영양기관 현황
노인코호트DB	2002년 자격유지자 중 만 60세 이상 대상자 중 약 55만 명 (모집단의 10%)	2002 ~ 2015년 (14개년)	2,749,045	사회·경제적 자격 변수(장애 및 사망 포함), 의료이용(진료 및 건강검진)현황, 영양기관 현황, 노인장기요양 서비스 현황
영유아검진코호트 DB	2007년 자격유지자 중 15~64세(생산가능인구) 직장 여성 약 18만 명(가입자 약 360만 명의 5% 무작위 추출)	2007 ~ 2015년 (9개년)	233,688	사회·경제적 자격 변수(장애 및 사업장 포함), 의료이용(진료 및 건강검진)현황, 영양기관 현황
직장여성코호트 DB	영유아검진 1~2차를 1회 이상 받은 전체 수검자 중 2008~2012년 출생자를 추출하여 각 출생연도별 5% 표본 추출	2008 ~ 2015년 (8개년)	368,226	사회·경제적 자격 변수(장애 및 사망), 의료이용(진료 및 건강검진)현황, 영양기관 현황

○ 코호트 정보에 대한 동의를 얻었는가?

○ 삭제청구권을 행사할 수 있는가?

2) 검토

○ 데이터 연계의 법적 근거가 있는가?

○ 개인정보보호원칙에 부합하게 수집되고 있는가?

○ 개인정보주체의 권리(삭제, 정정청구권 등)가 보장되고 있는가?

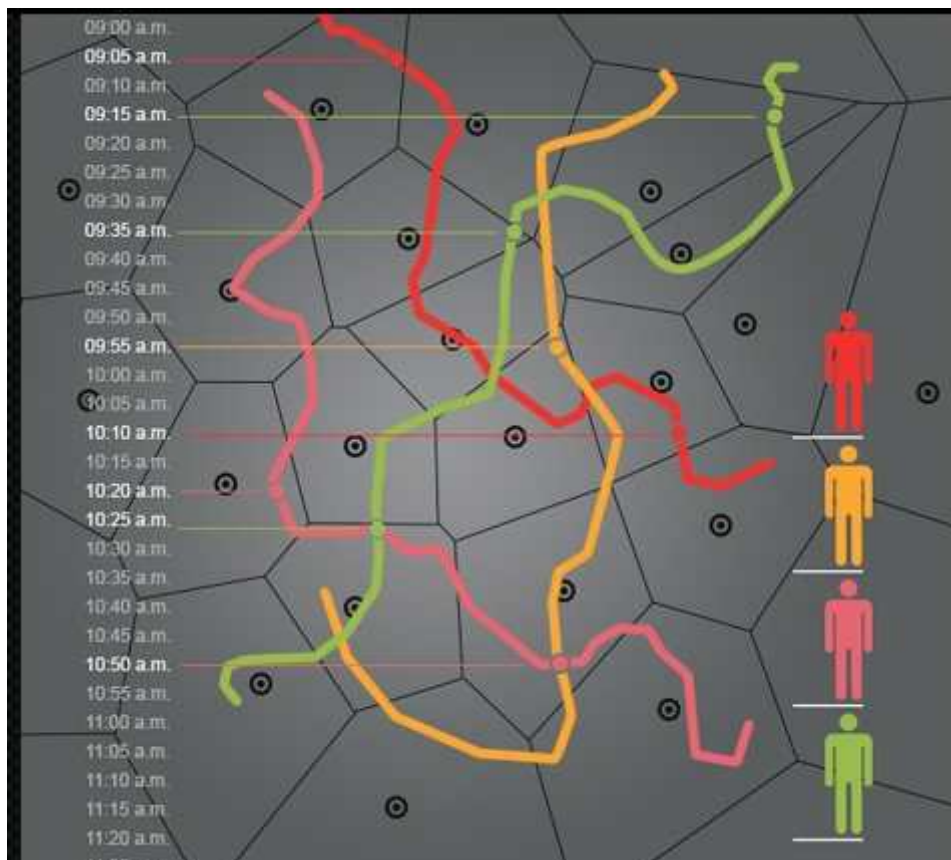
○ 연구나 공공목적의 활용에 대해서는 충분한 안전장치가 마련되어 있는가?

○ 절차가 투명하게 공개되어 있는가?

5. 익명 처리 또는 가명 처리 등의 처리와 재식별 가능성

1) 위치와 시간정보의 재식별 위험성

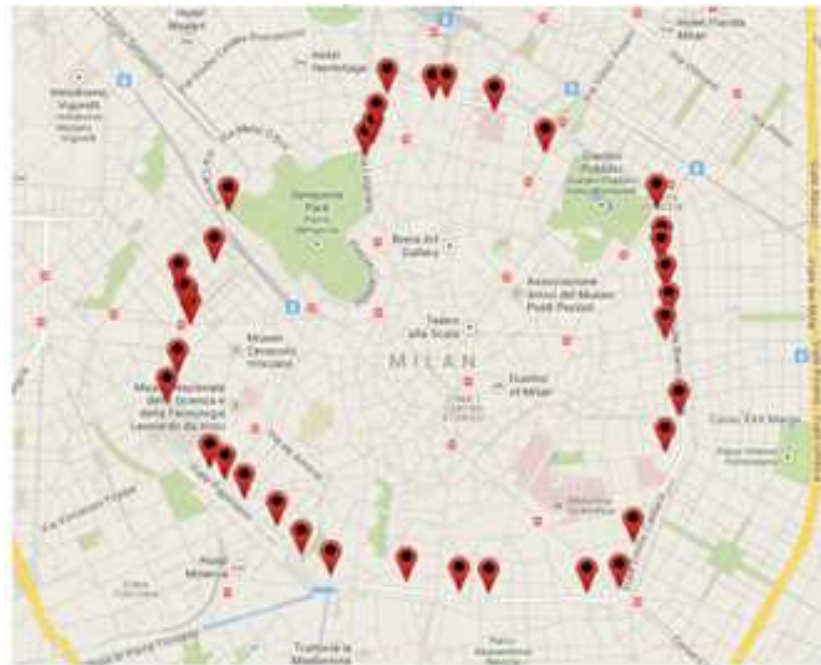
- 단 4조각의 위치와 시간 정보로 대부분의 정보가 재식별화 가능함
 - 위치와 시간정보는 재식별이 매우 용이함.
 - 특정 개인의 위치정보나 궤적정보는 익명화처리를 해도 외부 정보들을 활용하여 손쉽게 재식별이 가능.
 - MIT 미디어랩의 Yves-Alexandre de Montjoye 교수 등의 연구에 의하면 단 4조각의 외부 정보만 있어도 위치정보가 특정인을 재식별화할 수 있음이 확인됨 (Unique in the Crowd: The privacy bounds of human mobility 2013년).
 - 아래의 그림처럼 빨간색 이용자의 경우는 두 개의 시공간 정보, 녹색은 3개, 주황색은 2개, 분홍색은 2개의 시공간 정보로 각각 한 사람이 특정되는 것을 볼 수 있음.
 - 녹색과 분홍색, 주황색과 분홍색 이용자가 만났다는 것도 확인할 수 있음.



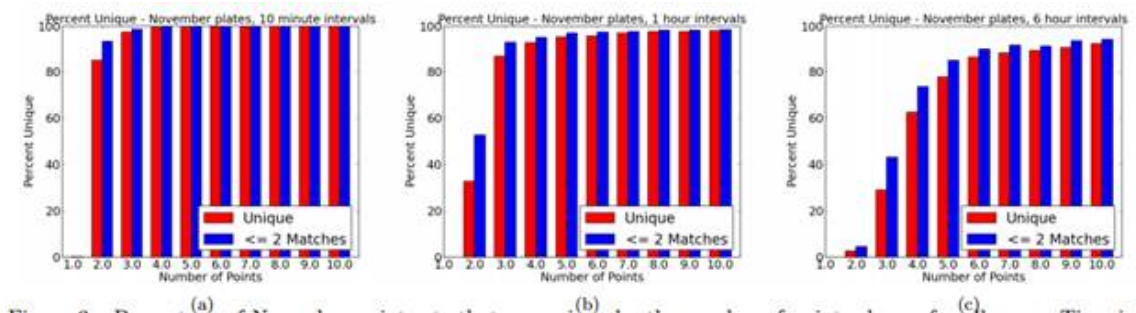
<그림2-1> 4조각의 위치와 시간으로 특정인 재식별 가능

- 밀라노 톨게이트 데이터 분석 결과

- 2013년 11월 밀라노 톨게이트 데이터(60만대) 분석 결과(Identifiability of Vehicle Tollgate Records: The Milan Tollgate Dataset(Nick Manfredi)
- 단, 2-3곳의 정보만으로도 개인 재식별 가능함.



<그림2-2> 밀라노시의 교통 게이트 위치



[11월에 단일한 사람이 식별되기 위해서 필요한 지점의 수. (a)는 10분 단위, (b)는 1 시간 단위, (c)는 6시간 단위]

2) 상품구매정보의 재식별 위험성 분석

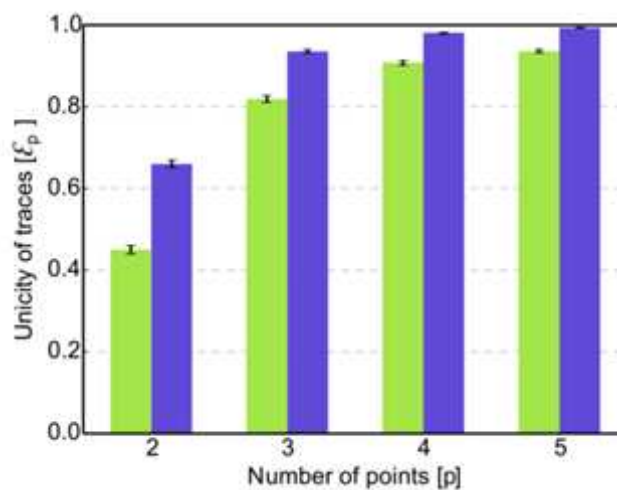
○ 단 몇 조각의 정보로 재식별 가능함

- 카드 구매내역 정보는 익명처리를 하더라도 재식별 가능성이 매우 높음.
- MIT 미디어랩의 Yves-Alexandre de Montjoye 교수의 연구

- 3개월의 110만 명의 신용카드 사용기록에 1만 개의 가맹점이 있는 경우, 단 4개의 시공간 점이면 개인을 고유하게 식별하는 비율이 90%에 달함.
- 페이스북이나 블로그, 그 밖의 다른 방법으로 4개의 시공간 점만 알면 익명화되어 있는 구매내역 정보의 더미에서 90%의 비율로 그 사람을 식별해 낼 수 있음.
- 시공간 점을 확인하는 방법은 소셜 미디어 정보 등 활용할 수 있는 정보가 무수하게 많음.



<그림2-3> 화살표는 익명화된 데이터에서 7abc1a23 이용자의 궤적을 보여줌



<그림2-4> 익명화된 구매내역 정보에서 시공간 정보로 재식별 가능 여부

- 익명화한 구매내역 정보에서 몇 개의 시공간 정보로 유일한 주체가 식별되는지 (재식별률)를 그래프로 표시한 것. 수평축은 시공간 정보의 수를 표시하고, 수직축은 유일한 주체가 식별되는 비율을 표시함. 보라색의 막대는 구매가격의 범위를 알 경우의 재식별률임. 구매내역정보는 3~4개의 시공간 정보만으로도 특정 개인이 재식별될 가능성이 8~90%를 넘는다는 것을 알 수 있음.
- 시공간점이 많아질수록 식별률은 높아짐.
- 구매자의 거래 가격을 알면 개인식별 가능성을 22% 증대시킴.
- 남성보다 여성은 재식별이 더 잘 되고, 소득수준이 높아도 재식별이 더 잘됨.
- 카드 구매내역 정보의 경우 아무리 익명화 기술을 활용하더라도 재식별 가능성은 매우 높다는 것임.

3) 그렇다면 건강정보는?

- 재식별가능성이 매우 높음.
- 위치, 시간 정보 포함되어 있음.
- 상병정보는 상품구매정보 보다 훨씬 더 정확하게 재식별이 가능함.
- 명세서당 여러 개의 질병코드를 기재함(종합병원의 경우 평균 3.07개).

	대분류 (장)	중분류 (항목군)	소분 (3단위분류)	세분 (4단위분류)	세세분류 (5단위분류)
KCD-7	22개	267개	2,082개	12,505개	6,746개

- 예시
- 100만 명의 환자 정보를 환자 이름이나 주민등록번호 등을 일련번호로 바꾸어 놓음.
- 상병정보 중 민감한 일부 정보는 일부만 표시함.
- 병원 이름을 일련번호로 치환함.

○ 청구정보에 포함된 정보들

명세서인번(이)불 (2019.05.30)

W.KJC20001 2007.05.30

요양급여비용명세서(의과)

제이 제호 전 수 심사부 명일련 00001

3-1 5888 UR2

610 03 정역,률 보험종별 공상등 선별구분 정결 접수일자 2007.09.20 접수번호 4039783 도학경로 입원경로

진료형태 의과외래 정률 건강보험

요양기관 명 형 기 호

중<보장기관>번호 72205494880

가입자 (세대주) 성 명 기 호

수진자 주민번호 <보장기관> 기 호

성 명 장 애 소영자료 마취전문동의서명

병 록 번 호

상 병 명 및 부 위 상병분류 기호 수술 과목 상해 특청기호 요양계시일

상세불명의 협심증 1209 내과 2007.02.28

상세불명의 만성 골관(신장)기능상실 NI89 내과 2007.02.28

초 회수 1 지 회수 2 입.내원일 1 내원일 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30

진 야.공 진 야.공 요양일수 1 내원일

약품두약금

CT 촬영

MRI 촬영

소계 I 21,068

II 14,160

가산 들 25%금 3,540

본연상한액

급여비출액 38,760

본인부담금 19,380

장 애 불

청 구 액 19,380

00/100본이

배금여출액

진료비출액

보훈청구액

기본적대

가산적대

일반 내역 - 수진자 인구 특성, 수진자 개인 ID, 요양기관기호, 서식코드, 진료과목코드, 입원외래구분, 의료급여종별코드

19

H I A 건강보험심사평가원 Health Insurance Review & Assessment Service

변수	속성명	자료형	크기
KEYCODE	명세서 연결코드	숫자	8
RECNO	요양일수	숫자	3
NO	수진자고유번호	숫자	8
YNO	요양기관 고유번호	숫자	8
dif	입원, 외래 샘플구분	문자	1
VSCN	입내원일수	숫자	3
EDEC_ADD_RT	심결가산율	숫자	3
DMD_TYPE_CD	청구형태	문자	1
FORM_CD	서식구분코드	문자	2
INSURER_CD	보험자코드	문자	1
MAIN_SICK	주상병	문자	5
SUB_SICK	부상병	문자	5
OPR_YN	수술여부	문자	1
DSBJT_CD	진료과목코드	문자	2
age	수진자연령	숫자	3
gen	성별구분	문자	1

변수	속성명	자료형	크기
IN_PAT_CORS_TYPE	입원도착경로구분	문자	2
OFFC_INJ_TYPE	공상구분	문자	1
DMD_TRAMT	청구요양급여비용총액	숫자	8
DMD_SBRDN_AMT	청구본인부담금	숫자	8
DMD_JBRDN_AMT	청구보험자부담금	숫자	8
DMD_DRG_NO	청구DRG번호	문자	10
MACL_CD	의료급여종별코드	문자	10
DMD_TYPE	청구구분	문자	1
EDEC_SBRDN_AMT	심결본인부담금	숫자	8
RECU_FR_DT	요양개시일자	날짜	8(YMMMDD6.)
FST_IN_PAT_DT	최초입원일	날짜	8(YMMMDD6.)
SPC_SYM_TYPE	특정기호구분	문자	4
DIAG_RSLT_TYPE	진료결과구분	문자	1
INJ_OUT_PRY_FACTOR	상해외인구분	문자	1

DGSBJT_CD (진료과목코드)

- 요양기관에서 청구 시 기재한 진료과목 코드

[의과]

코딩	설명	비율(%)	코딩	설명	비율(%)
00	일반의과	2.35	14	피부과	4.11
01	내과	24.71	15	비뇨기과	2.19
02	신경과	1.13	16	영상의학과	0.19
03	정신건강의학과	1.43	17	방사선종양학과	0.05
04	외과	3.11	18	병리과	0.00
05	정형외과	14.19	19	진단검사의학과	0.02
06	신경외과	1.44	20	결핵과	0.01
07	흉부외과	0.24	21	재활의학과	0.59
08	성형외과	0.09	22	핵의학과	0.01
09	마취통증의학과	0.92	23	가정의학과	2.43
10	산부인과	2.78	24	응급의학과	0.61
11	소아청소년과	7.52	25	직업환경의학과	0.02
12	안과	5.16	26	예방의학과	0.01
13	이비인후과	5.82	99	해당사항 없음	0.00

● DGSBJT_CD (진료과목코드)

• 요양기관에서 청구 시 기재한 진료과목 코드

[의과]

코딩	설명	비율(%)	코딩	설명	비율(%)
00	일반의과	2.35	14	피부과	4.11
01	내과	24.71	15	비뇨기과	2.19
02	신경과	1.13	16	영상의학과	0.19
03	정신건강의학과	1.43	17	방사선종양학과	0.05
04	외과	3.11	18	병리과	0.00
05	정형외과	14.19	19	진단검사의학과	0.02
06	신경외과	1.44	20	결핵과	0.01
07	흉부외과	0.24	21	재활의학과	0.59
08	성형외과	0.09	22	핵의학과	0.01
09	마취통증의학과	0.92	23	가정의학과	2.43
10	산부인과	2.78	24	응급의학과	0.61
11	소아청소년과	7.52	25	직업환경의학과	0.02
12	안과	5.16	26	예방의학과	0.01
13	이비인후과	5.82	99	해당사항 없음	0.00

명세서일련번호 (20table)		구분	코드	분류명 및 품명	단가	일투	총투	금액	비고	순번
				01항.진찰료						
		1	AA256	재진진찰료-종합병원 종합전문요양 기관에 설치된 경우를 제외한 치료 대학부속치과병원	10,410.0	2.00	1	20,820		1
				01 / A / 1	10,650.0	2.00	1	20,620		2
				03항.투약및처방전료						
		3	A09700641	117메니폰정 10mg/A 107501ATB	1.0	2.00	62	124	2007.02.28 여외 [99]	3
		3	A05701481	117메트라닐정 25mg/A 107502ATB	2.0	1.00	62	124	2007.02.28 여외 [99]	4
				09항.검사료						
		1	C3793	전해질(염소)Cl	1,210.0	1.00	1	1,210 +		5
		1	C3795	전해질(총칼슘)TotalCa	1,210.0	1.00	1	1,210 +		6
		1	C3730	요소질소(NPN포함)	1,640.0	1.00	1	1,640 +		7
		1	C3750	크레아티닌	1,370.0	1.00	1	1,370 +		8
		1	C3791	전해질(소듐)Na	1,210.0	1.00	1	1,210 +		9
		1	C3792	전해질(포타슘)K	1,210.0	1.00	1	1,210 +		10
		1	C3710	당검사(반경량)	870.0	1.00	1	870 +		11
		1	B1010	혈색소(광전비색)	1,050.0	1.00	1	1,050 +		12
		1	B1020	헤마토크리트	680.0	1.00	1	680 +		13
		1	B1040	적혈구수	680.0	1.00	1	680 +		14
		1	B1050	백혈구수	680.0	1.00	1	680 +		15
		1	B1060	혈소판수	800.0	1.00	1	800 +		16
		1	B1091	백혈구백분율(혈액)	1,610.0	1.00	1	1,610 +		17

처치, 수술, 원내 약 처방 청구금액 등 상세 의료 서비스 정보

○ 재식별 가능성

- 보험청구를 받은 보험회사에서: 보험회사는 보험청구를 받은 환자의 진료정보에서 언제(시간), 어떤 상병으로, 어떤 진료를 받았는지를 확인함. 해당 자료로 매칭 할 경우, 재식별 가능함.

일시	요양기관	상병	진료 등
2014. 5. 1.	A병원	바이러스성 장염	주사 약물 처방
2014. 7. 1~5	B병원	급성 C형 간염	주사, 약물, 검사, 입원
2014. 8. 1.	C병원	상세불명의 협심증	약물 처방
2014. 10. 1.	A병원	상세불명의 만성 콩팥 기능상실	약물 처방

- 병원에서: 병원이 보유하고 있는 의료정보로 개인 재식별 가능
- 신문기사, 소셜미디어 정보 등에서: 입원, 교통사고 등의 기사나 소셜미디어 정보를 통해서 재식별 가능함. Latanya Sweeney(Harvard University)는 신문 기사를 통해서 익명화 처리를 한 81명의 의료기록에 대해 35명을 재식별함(Matching Known Patients to Health Records in Washington State Data).

How Someone Can Re-identify Your Medical Records

Even though patient names aren't attached to the health records being sold to marketers, an anonymous hospital record can be cross-referenced with publicly available information to "re-identify" a person. Bloomberg News, in conjunction with Latanya Sweeney, director of Harvard University's Data Privacy Lab, re-identified 35 people out of 81 cases searched in a database of hospital discharge records that Washington State sold to the public for \$50. One patient agreed to let Bloomberg use his case to show how re-identification can work:

GRAPHIC: DAVE MERRILL / BLOOMBERG VISUAL DATA

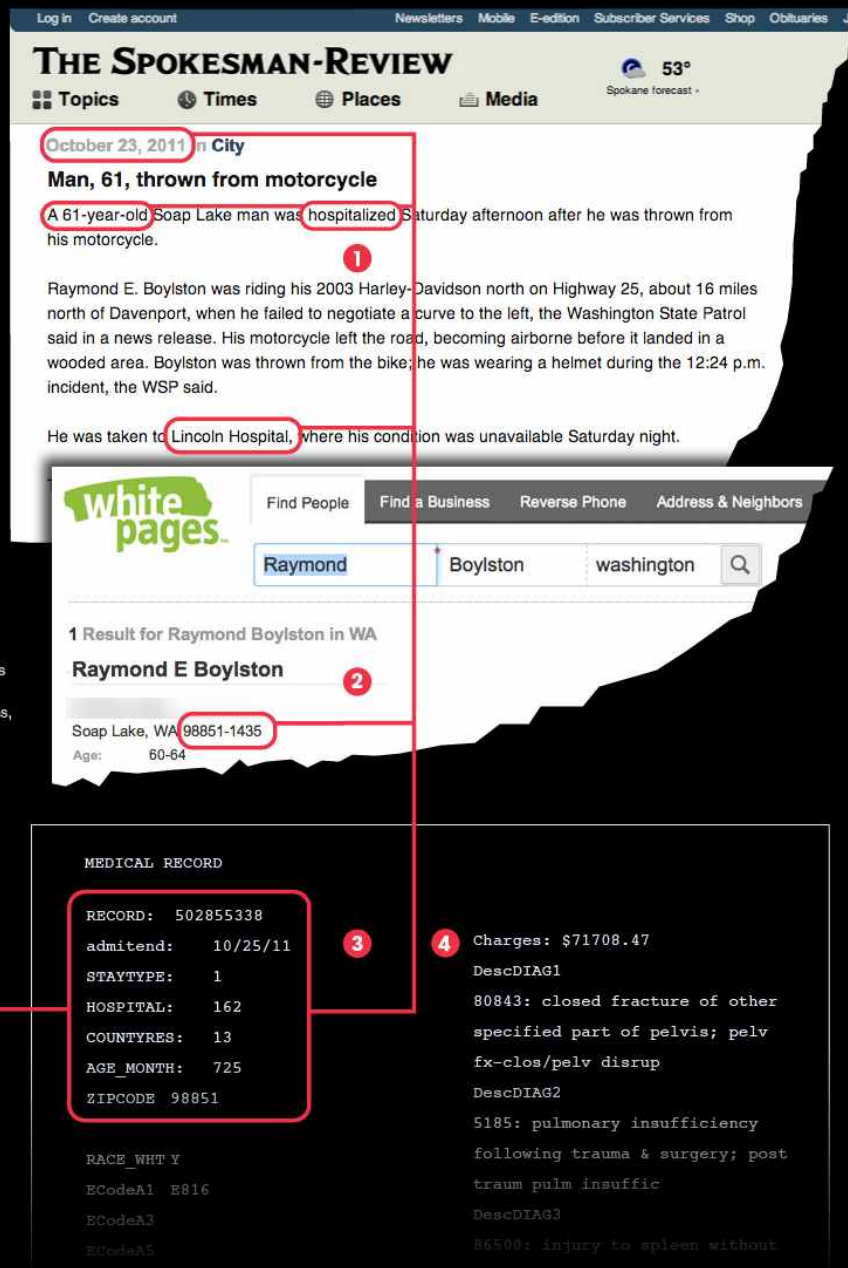
1 Washington state news articles were searched for the word "hospitalized." Most of these articles included the person's name, age, town of residence and reason for hospitalization.

2 The person's name, age and residence is searched online. Several online sites will reveal ZIP codes associated with the search terms.

3 Taking the newly learned ZIP code, plus the patient's age¹, approximate date and location of hospitalization², a match can be found in the health record dataset purchased from the state.

4 Within the electronic record is private patient information including: physician diagnoses, procedures and payment information.

5 
Raymond E. Boylston, the man identified in the news article, is linked to "anonymous record" #502855338.



THE SPOKESMAN-REVIEW
October 23, 2011 in City
Man, 61, thrown from motorcycle
A 61-year-old Soap Lake man was hospitalized Saturday afternoon after he was thrown from his motorcycle.
Raymond E. Boylston was riding his 2003 Harley-Davidson north on Highway 25, about 16 miles north of Davenport, when he failed to negotiate a curve to the left, the Washington State Patrol said in a news release. His motorcycle left the road, becoming airborne before it landed in a wooded area. Boylston was thrown from the bike; he was wearing a helmet during the 12:24 p.m. incident, the WSP said.
He was taken to Lincoln Hospital, where his condition was unavailable Saturday night.

white pages
Find People Find a Business Reverse Phone Address & Neighbors
Raymond Boylston washington
1 Result for Raymond Boylston in WA
Raymond E Boylston
Soap Lake, WA 98851-1435
Age: 60-64

MEDICAL RECORD
RECORD: 502855338
admitend: 10/25/11
STATTYPE: 1
HOSPITAL: 162
COUNTYRES: 13
AGE_MONTH: 725
ZIPCODE 98851
RACE_WHT Y
ECodeA1 E816
ECodeA3
ECodeA5
Charges: \$71708.47
DescDIAG1
80843: closed fracture of other specified part of pelvis; pelv fx-clos/pelv disrupt
DescDIAG2
5185: pulmonary insufficiency following trauma & surgery; post traum pulm insuffic
DescDIAG3
86500: injury to spleen without

¹ - Boylston was 60 years old at the time of hospitalization (which the discharge record shows accurately - 725 months). Boylston's age was inaccurately reported in the news article.
² - Boylston was transported to Lincoln Hospital. He was treated at Providence Sacred Heart in Spokane.

Sources: theDataMap.org, Bloomberg research. Graphic by Dave Merrill / Bloomberg Visual Data.

 Bloomberg Visual Data

6. 코호트 DB의 제공, 표본 데이터셋의 제공

1) 코호트 DB의 제공

- 코호트 DB를 아래의 비식별화를 한 후 제공



- 재식별 가능성은 매우 높음
 - 요양기관 대체식별번호는 충분히 재식별이 가능함.
 - 민감상병 외의 상병으로 충분히 재식별 가능함.

2) 표본 데이터셋 제공

- 민간보험사 제공
 - 건강보험심사평가원에서 제출한 자료에 따르면, 건강보험심사평가원은 2014년 7월부터 2017년 8월까지 KB생명보험 등 8개 민간보험사 및 2개 민간보험연구기관이 보험료 산출과 보험상품개발 등을 위해 요청한 '표본 데이터셋'을 1건당 30만 원의 수수료를 받고 총52건(약 6,420만 명분) 제공.

[표-1] 민간보험사 및 민간보험연구기관 빅데이터(표본) 제공 현황 (단위: 건)

구분	제공받은 기관(제공된 데이터건수)	제공받은 데이터	활용목적
민간보험사	AIA생명(12), KB생명(10), KB손해(3), SCOR(2), 롯데손해(2), 미래에셋생명(3), 현대라이프생명(6), 흥국화재해상보험(9)	각 연도별 전체/입원/고령/소아 환자데이터셋 등	-진료환자 분석 -보험상품연구개발 위한 위험률 산출 등
민간보험연구기관	보험개발원(4), 보험연구원(1)		-순보험료 산출 -위험률 산출 등

※데이터 요청 1건 당 일괄적으로 30만원의 수수료 받음. ※건강보험심사평가원 제출자료

- 민간보험사 등이 받아간‘표본 데이터셋’은 모집단의 특성을 잘 대표할 수 있는 표본을 추출하여 구성한 자료.
 - 대상은 전체(140만 명)/입원(110만 명)/고령(100만 명)/소아청소년(110만 명)환자로 구분.
 - 성별, 연령 등을 담은 일반내역 뿐 아니라 진료행위 등을 담은 상병내역과 주상병 등이 담긴 진료내역, 원외처방내역으로 구성.

[표-2] 건강보험심사평가원의 표본 데이터셋 구성

구분	구 분	산출 기준
대상 인원	전체 환자	연 단위 전체 환자 약 140만명 (모집단의 3%)
	입원 환자	연 단위 입원환자 약 70만명 (모집단의 13%)
	고령 환자	연 단위 외래환자 약 40만명 (모집단의 1%)
	소아·청소년 환자	연 단위 65세 이상 환자 약 100만명 (모집단의 20%)
자료 구성	일반내역	연 단위 20세 미만 환자 약 110만명 (모집단의 10%)
	상병내역	성별, 연령, 연령군, 보험자코드, 의료급여종별코드, 요양급여비용총액 정보 등
	진료내역	외래 또는 입원 시 발생하는 진료행위(행위별수가), 의약품(주된 성분만 제공), 실시횟수 정보 등
	원외처방내역	명세서에 기재되어 있는 주상병·부상병에 대한 정보(KCD코드, 진료과목) 등
		외래환자 원외처방으로 이루어지는 의약품(주된 성분만 제공)에 대한 처방전 정보 등

※건강보험심사평가원 제출자료

○ 제공 현황

- 보험개발원: 2009~2011년 입원환자데이터셋, 2013년, 2015년 전체환자데이터셋
- KB생명: 2009 ~ 2015년 입원환자데이터셋, 2013 ~ 2015년 전체환자데이터셋
- KB손해보험: 2011 ~ 2013년 입원환자데이터셋
- 흥국화재해상보험: 2009~ 2011년 입원환자데이터셋, 고령환자 데이터셋, 소아청소년데이터셋, 전체환자데이터셋
- 롯데손해보험: 2012, 2013년 전체환자데이터셋

- 미래에셋생명: 2013 ~ 2015년 입원환자데이터셋
- SCOR 재보험사: 2014, 2015년 입원환자데이터셋
- AIA 생명: 2010 ~ 2015년 입원환자, 전체환자 데이터셋
- 현대라이프생명: 2010 ~ 2015년 전체환자데이터셋

○ 소위 비식별 조치

○ 원천DB에서 환자이름 삭제, 주민번호를 H-PIN(대체키)으로 변환



< 2차 비식별화 조치 방법 >

처리기법	조치 내용
가명처리	· 1차 비식별화 대체키에 임의의 난수 부여 후 1차 비식별자 삭제 · 요양기관 기호 및 명세서 번호 임의의 난수 부여
데이터 삭제	· 요양기관 주소는 시·도까지 제공 · 법정 전염병 및 100세 이상 환자 삭제 · 제약업체 영업정보 보호를 위하여 의약품코드 삭제 · 의약품 주성분 중 제형정보 일부 삭제
데이터 마스킹	· 치료재료업체의 영업정보 보호를 위하여 업체를 식별할 수 있는 세분류코드 마스킹

○ 재식별 가능성

- 몇 개의 환자정보만으로도 충분히 재식별 가능함.

○ 개인정보보호법 위반: 환자의 동의 없는 민감한 건강정보 제공에 해당함.

7. 결론

- 현재 국민건강보험공단, 심평원, 질병관리본부 등의 건강정보 보유, 활용은 많은 문제가 있음.
- 정책 목적, 연구 목적의 활용에 대해서도 충분한 안전장치를 마련해야 함.
- 그와 같은 안전장치가 마련되어 있지 않기 때문에 정책 목적, 연구 목적의 활용에 대한 신뢰가 없음.
- 이와 같은 문제를 해결하고, 신뢰를 얻기 전에 조금하게 빅데이터 활용을 추진하는 것은 문제를 야기할 수 있음.

보건의료 빅데이터 산업화의 몇 가지 문제들

변혜진 / 연구공동체 건강과대안 상임연구원

- 박근혜 정부 시기부터 거침없이 진행된 빅데이터 산업화 논리가 문제인 정부로 이어져 ‘보건의료 빅데이터 플랫폼 구축’ 예산 심의 및 복지부 ‘보건의료 추진 전략’ 발표 등을 앞두고 있음. 그러나 박근혜 정부의 빅데이터 산업화 추진은 그 내용과 필요, 정보 제공자에 대한 충분한 사전 설명과 동의 절차가 없는 전형적인 묻지마식 행정이었으며, 기업과 부패 권력의 합작품이기도 한 ‘창조경제’의 키워드이자 규제 완화를 위한 마케팅 도구로 활용된 측면이 큼.
- 이런 점에서 볼 때 2016년 다보스 포럼 이후 ‘4차 산업혁명’으로 간판을 바꾼 창조경제에 기초를 둔 박근혜 정부식 빅데이터 산업화 방안은 전면 재고될 필요가 있으며, 이를 기반으로 혈세가 집행되고 예산이 기획된 공적 재정 투자와 이와 관련해 규제 완화가 전제된 법 제도들은 재논의와 검토가 필요함.
- 발제에서도 지적되었듯이, 보건의료 빅데이터 산업화 전략으로 집행된 ‘정밀의료’ 사업이나 ‘보건의료 빅데이터 플랫폼 구축’ 사업은 그 실효성에 대한 사회적이고 과학적인 재검토가 필요함. 또한 현행법에서 보장하고 있는 개인정보 보호에 대한 법률과 환자 정보를 보호하기 위한 의료법 등과의 충돌 등에 대한 법제도적 검토가 필요함. 복지부는 이전 정부에서 추진돼 오던 맹목적인 보건의료 빅데이터 산업화에 대한 시민사회의 우려에 대해 주의를 기울일 필요가 있음. 현재 복지부의 추진 전략은 기업 민원을 해결해주려는 개인 의료/건강 정보 민영화 전략과 다름

바가 없음. 공공의 목적을 위한 빅데이터 활용은 사실상 이를 위한 레토릭에 불과함. ‘갈등과 논란’ 있는 사안일 수밖에 없는 문제를 묻지마 행정으로 해결해서는 안됨. 이에 복지부의 추진 전략이 가진 몇 가지 문제를 지적하고자 함.

1. 개인 진료 기록 및 건강 데이터의 특수성 간과

- 정부가 추진하는 빅데이터 사업 일반과 보건의료 분야에 대한 논의는 그 틀을 달리해야 함. 누구나 인정하듯이 정부가 이용하겠다고 하는 보건의료 데이터의 경우 개인정보 중에서 가장 예민하고 민감한 정보에 해당되기 때문임. 데이터의 성격상, 정보 제공자가 매우 취약한 상태에서 제공한 정보가 포함돼 있으며 (몸이 아파 치료를 구하는 개인정보제공 동의서라는 의미), 이러한 정보는 동의된 상태에서 활용 목적상 개인의 과거의 삶이 연결된 ‘기록’ 이자 가족 등의 부가적 정보가 포함된 정보(개인의 질병 상태는 과거로부터 연결된 결과이며, 의료인은 이를 치료하기 위해 개인의 데이터와 기록을 묻고 확인하고, 이를 다시 심평원이 심사하는 진료기록으로 남길 수 있음) 이기 때문임.
- 따라서 매우 신중하고 예민하게 보호될 필요가 있으며, 공공의 목적으로 사용될 시에도 개인의 정보 이용 동의서가 반드시 필요한 데이터 임. 현재 공공기관에 수집된 국민의 개인 의료 기록들은 단순한 수치가 아니라 한 개인의 과거와 현재가 연결돼 있는 기록이라는 점을 간과해서는 안됨. 이러한 점에서 관련 데이터가 유출 되었을 시 그 개인에게 돌아가게 될 심각한 피해는 사실상 상상할 수 없는 고통을 안겨줄 수 있으며, 주민번호제도라는 특이성이 있는 우리나라의 현실에서 그 기록을 지우기도 어렵다는 점에서 매우 신중해야 할 문제임.
- OECD 이사회조차도 17년 1월 ‘보건의료 데이터 거버넌스에 대한 권고 (Recommendation of OECD Council on Health Data Governance)’안을 내놓고, “건강관련 데이터는 본질적으로 민감하고, 데이터 공유와 사용 확대는 데이터의 손실 위험을 증가시키거나 유출 및 오남용 위험을 야기해서 개인에게 개인적, 사회적, 재정적인 피해를 끼칠 수 있고, 보건의료서비스 제공자와 정부에 대한 대중의 신뢰를 떨어뜨릴 수 있다”¹⁾고 지적한 바 있음. 각국 보건 장관 회의에서 채

1)

<http://www.oecd.org/health/health-systems/Recommendation-of-OECD-Council-on-Health-Data-Governance-Booklet.pdf>

택된 이 권고 5장에서는 정보 이용에 대한 개인 정보 제공자에게 충분한 설명 후 동의를 구하는 (informed consent) 적절한 절차도 제시하고 있음. “a) 개인의 동의가 필요한 경우, 그 결정을 위해 사용되는 기준이 명확해야 하고, 유효한 동의를 구성하는 요소가 무엇이고, 어떻게 동의를 철회할 수 있는지도 명확해야 함. 동의를 구하는 것이 불가능하거나 비현실적이거나 건강 관련 공공의 목적과 부합하지 않을 경우, 동의를 대신할 수 있는 적절한 대안 및 예외가 명확해야 하고, 이러한 과정에 부합하는 보호조치가 뒤따라야 함. b) 개인 건강정보 처리가 동의에 기반한 경우, 이 동의는 충분한 설명이 자유롭게 제공되는 경우에만 유효 함. 개인이 장애의 정보 사용에 대해 동의하거나 철회할 때 그 방법이 명확하고 눈에 잘 띄고 사용하기 쉬웠을 때 유효“ 하다고 명시함.

Consent mechanisms should provide:

- a. Clarity on whether individual consent to the processing of their personal health data is required, and, if so, the criteria used to make this determination; what constitutes valid consent and how consent can be withdrawn; and lawful alternatives and exemptions to requiring consent, including in circumstances where obtaining consent is impossible, impracticable or incompatible with the achievement of the health-related public interest purpose, and the processing is subject to safeguards consistent with this Recommendation.
- b. That, where the processing of personal health data is based on consent, such consent should only be valid if it is informed and freely given, and if individuals are provided with clear, conspicuous and easy to use mechanisms to provide or withdraw consent for the future use of the data.

- 그러나 정부는 이런 개인 민감 정보에 대한 국민의 동의 절차를 생략한 채 박근혜 정부 하에 진행되던 추진 방식과 전략을 그대로 답습하고 법을 우회하는 방식으로 추진하려 함. 이는 큰 사회적 혼란과 불신을 초래할 것이며, 의료인-환자-건강보험에 대한 불신을 야기하는 단초가 될 것임.
- 현재 정부가 내 놓고 있는 보건의료 빅데이터 추진 전략의 핵심은 두 가지임. 1) 보건의료 데이터 간 ‘연계 활용’ 2) 공공 외 민간 기관 (기업) 데이터 연계 및 상업적 제공/ 활용 허용. 그러나 세계 어느 나라도 민간과 공공 정보의 구분 없이 개인 건강 정보를 이렇게 연계 활용하도록 한 나라가 없음. 공공 기관 데이터를 ‘연계 활용’ 하도록 허용하는 일부 국가들의 경우에는 그 데이터 활용이 공공의 목적에 부합하며, 해당 연구윤리 심사가 완료된 후, 공적 목적에 한하여 허용

하고 있으며, 연구보고서는 그 목적 심사에 타당하게도 모두 공개되는 원칙이 적용됨. 즉 국가가 공공 데이터, 국민의 개개인 진료 건강 정보를 연계 제공하는 목적이 매우 분명한 공익적 목적에 한해 있다는 것이며, 그 결과가 모두에게 제공되며, 공공적 통제와 관리가 가능하다는 것임. 그러나 지금 현 정부가 추진하고 있는 정책은 공적으로 모은 개인 진료/건강정보에 민간기업이 모은 데이터를 ‘연계’ 해 이를 ‘상업적으로 활용’ 하는 것을 허용하겠다는 것임. 이는 기업의 상품 개발이나 마케팅을 위해 공공 데이터를 제공할겠다는 것 그 이상도 이하도 아니며, 연구를 위한 이용이라 한정한다해도 그 과정에 대한 공공적 통제 방안과 개입의 제도가 전무하며, 이후 보고서는 ‘영업비밀’이라는 이유로 공개되지 않을 것임.

- 이는 기업이 참여하는 대학 산학협력단 차원의 연구가 ‘공적 연구’ 목적으로 허가를 받아도 그 결과에 대해서는 영업비밀로 제한되고 있는 현실을 보아도 잘 알 수 있으며, 결국 기업의 상업적 이용이 허용되는 범주에서 진행되는 보건의료 빅데이터 사업은 그 연구의 내용과 과정 그리고 그 결과의 공개 및 투명성을 보장할 수 없음. 이는 정보 제공자의 정보의 보안과 안전이 지켜질 의무의 주체가 불분명해진다는 것을 의미함.

2. ‘건강 정보’ 에 대한 정의와 법적 보호 조치가 우선

- 복지부는 사실상 2016년 자신들이 로펌에 의뢰해 받은 ‘보건의료 빅데이터 플랫폼 구축’ 사업이 불법이라는 자체 보고서도 무시하고 있는 지경임.
- 이러한 복지부의 성급한 추진은 현재 법으로 보호하고 있는 개인정보 정의와 범위를 기업 맘대로 해석하게 하는 문제를 양산하고 있음. 2015 SK텔레콤 전자의 무기록 수입 사태나, 약학정보원과 지누스 등이 IMS헬스에 국민 대다수의 개인질병정보 판매한 사태나, 현 심평원 사태에 이르기까지 국가기관인 검찰과 각 부처가 서로 다른 법리 해석을 하며, 국민의 이익보다는 기업의 편익에 치중해 법을 규제 밖에서 해석하는 논리에 근거를 대고 있음.
- 정부가 추진하겠다는 전략은 현재는 텍스트 데이터라 볼 수 있는 건강보험공단, 심사평가원, 질병관리본부, 유전자검사 결과 등의 연계이지만, 보건의료 플랫폼 구축을 통해 여기에 민간기관의 데이터가 연계되거나 공공데이터의 민간 기관 결

합이 허용되면, 현재 모바일 앱이나 IOT 등에 음성적 형태로 수집되고 있는 개인 신체·건강·생활정보의 집적이 합법화되고 어떤 방식으로 활용될지 예측할 수 없는 상황에 놓이게 됨. 따라서 정부는 이러한 위험을 먼저 인식하고 이를 보호할 조치를 간구하는 것이 우선이어야 함.

- 유럽 등 국가에서 이러한 문제로 인해 ‘개인 건강정보’에 대한 보다 엄격한 정의를 논의하고 이를 법 제도화로 보호하는 절차를 진행하고 있음. 따라서 복지부가 주무부처로서 제대로 된 자기 역할을 되찾는 것은 이러한 문제를 다른 행정 기관에 알리고 이를 위한 법 제도 보완 정책을 만들어 가는 것임. 박근혜 행정부에서 발표된 행안부의 ‘비식별화 가이드라인’을 기초로 민감정보에 해당하는 국민 개인 질병/건강정보와 음성적으로 수집 이용되고 있는 각종 신체·건강 정보에 매칭시키는 일의 합법화는 국민 건강과 복지를 책임진다는 주무부처가 나서서 할 일이 아님.
- 예를 들어, (가칭)‘건강정보 및 유전정보 처리 및 보호에 관한 특례법’ 제정 논의를 먼저 시작할 수 있음. 아래 내용은 유럽 등에서 제정되고 있는 내용을 근거로 개인 건강정보의 정의와 범위를 법제도 보호하기 위한 예임. 현재 무작위 수집되고 있는 개인 신체·건강 정보 데이터를 어떻게 관리하고 규제할 것인지에 대한 별도의 충분한 논의와 법제도가 마련된 후에야 현재 정부가 추진하려는 ‘보건의료 빅데이터’ 추진 전략에 대한 불신이 해소될 수 있을 것임.

제2조(정의) 이 법에서 사용하는 용어의 뜻은 다음과 같다.

1. “건강정보”란 살아 있는 개인의 육체적, 정신적 건강에 대한 정보로서, 「보건의료기본법」 제3조제2호에 의한 보건의료서비스 제공 과정에서 수집되는 정보를 포함하여, 과거, 현재, 미래의 개인의 건강 상태를 드러내는 정보를 말한다.
2. “유전정보”란 살아 있는 개인의 물려받거나 획득한 유전적 특성과 관련된 정보로서, 인체로부터 수집하거나 채취한 조직·세포·혈액·체액 등 인체 구성물 또는 이들로부터 분리된 혈청, 혈장, 염색체, DNA(Deoxyribonucleic acid), RNA(Ribonucleic acid), 단백질 등을 분석하여 얻은 개인의 생리, 건강에 대한 특수한 정보를 말한다.
3. “처리”란 건강정보 및 유전정보의 수집, 생성, 연계, 연동, 기록, 저장, 보유, 가공, 편집, 검색, 출력, 정정(訂正), 복구, 이용, 제공, 공개, 파기(破棄), 그 밖에 이와 유사한 행위를 말한다.
4. “정보주체”란 처리되는 정보에 의하여 알아볼 수 있는 사람으로서 그 정보의 주체가 되는 사람을 말한다.
5. “건강정보처리자”란 업무를 목적으로 건강정보 및 유전정보파일을 운용하기 위하여 스스로 또는 다른 사람을 통하여 건강정보 및 유전정보를 처리하는 공공기관, 법인, 단체 및 개인 등을 말한다.

다.

제3조(건강정보의 범위) 제2조에 의한 건강정보는 의료인, 보건의료인, 의료기관, 의료기기, 개인용 건강관리제품, 실험실 진단기기 등 정보를 수집하는 주체, 수집되는 방식 등과 상관없이 다음을 포함하는 개인정보를 말한다.

1. 「보건의료기본법」 제3조제1호에 의한 보건의료를 목적으로 개인을 고유하게 구별하고 식별하기 위해 부여된 숫자, 문자, 부호 등
2. 인체로부터 수집하거나 채취한 조직·세포·혈액·체액 등 인체 구성물 또는 이들로부터 분리된 혈청, 혈장, 염색체, DNA(Deoxyribonucleic acid), RNA(Ribonucleic acid), 단백질 등을 조사, 분석, 검사함으로써 생성된 정보
3. 정보주체의 질병, 장애, 질병 위험, 과거 병력, 치료, 생리학적 상태, 생의학적 상태 등에 대한 모든 정보

3. 차별과 배제를 위한 알고리즘

- 시민사회단체 용으로 회람된 보건의료 빅데이터 추진 전략의 경우, 사회 취약층에 대한 ‘관리’와 ‘통제’ 측면에서 데이터셋이 구성될 위험이 크다고 판단함. 보건복지 업무의 효율화를 이룰 수 있을지는 몰라도 이러한 데이터셋 구축은 한 사람의 건강을 구성하는 사회 결정 요인에 따를 때 특정 계층의 데이터셋, 혹은 ‘새로운 건강 블랙리스트’가 될 우려가 큼.
- 또한 인터넷 등 정보 접근권에서 불평등이 큼. 2016년 인터넷 이용 관련 정부 통계자료를 보면 전체 인터넷 이용률은 88.3%에 해당되지만, 계층별, 소득별, 지역별 인터넷이용률은 상당한 차이를 보임.²⁾ 무학의 경우 14.1%, 초등졸 이하는 65.3%로 줄어들고, 무직의 경우 53.5%, 농어업직은 67.1%로 떨어짐. ‘정보 격차(digital divide)’는 보건복지 이용에 장벽으로 작용하는 경우도 많음. 정보기술에 일방적으로 기댄 정부 정책은 이러한 격차를 줄이기보다는 애초에 배제시키는 결과를 초래할 수 있음.

2) 한국인터넷진흥원, 개인인터넷이용 통계자료 <https://isis.kisa.or.kr/statistics/?pageId=020200>

표 1 디지털정보화 수준		(단위: %)	
구분	2014년	2015년	2016년
장애인	60.2	62.5	65.4
저소득층	72.5	74.5	77.3
농어민	51.4	55.2	61.1
장노년층*	42.4	45.6	54.0
평균	50.1	52.4	58.6

* 일반국민의 디지털정보화 수준을 100으로 할 때, 일반국민 대비 4대 계층의 디지털정보화 수준을 의미

* '16년부터 장노년층 기준연령을 만50에서 만55세 이상으로 상향 조정

* 장노년층 기준연령 만55세 적용 사유

① 법률「고령자 고용차별금지 및 고령자고용촉진에 관한 법률」에서 고령자를 55세 이상으로 규정(동법 시행령 제2조)

② 고령층 정보화 교육 대상이 만55세 이상으로 정책 집행에 따른 성과를 동일한 기준으로 산출하기 위함

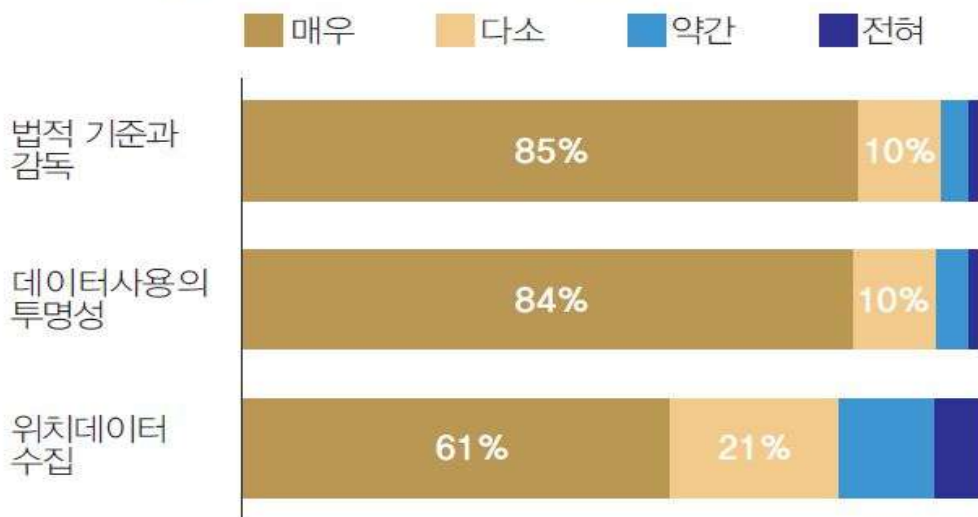
* '14~'15년 장노년층 디지털정보화 수준은 만55세 기준으로 재환산하여 산출

<그림3-1> '2016 디지털정보격차 실태조사', 미래과학부, 한국정보화진흥원, 2016



- 디지털 격차의 문제는 과학이 발전하면 모든 이들에게 혜택이 돌아갈 것이라는 보편주의 가정이 틀렸다는 것을 의미하기도 함. 과학기술은 어떤 맥락에서 어떻게 사회적으로 실행되는가에 따라 계층차별화가 있으며, 사회적 불평등을 재생산하는 도구로 기능하기도 함. 정부가 추진하는 빅데이터 산업화가 정보기술을 기반으로, 생명공학과 인공지능 등을 ‘융합’ 시키는 차원에서 홍보를 하고 있다는 점에서 디지털 격차에 대한 신중한 고려가 필요함. 결과적으로 사회 취약계층의 경우 입력할 데이터가 아예 비어 있는 경우도 있으며, 과잉돼 있는 경우도 있을 수 있음.
- 빅데이터가 차별과 배제의 알고리즘의 위험에 노출돼 있다는 지적은 여러 곳에서 제기되고 있음. 미국 백악관 보고서는 빅데이터의 활용 기회를 포착해야 하지만, 빅데이터 도구가 개인의 사적인 상세정보를 노출할 수 있음을 경고하고 이를 보완할 법 제도적 조치를 우선할 것을 권고한 바 있음.³⁾

그림 1 : 빅데이터 활용에 대한 미국 국민들의 우려



자료: Executive Office of the President(2014 May), *Big Data: Seizing Opportunities, Preserving Values*

<그림3-2> [미국] 알고리즘 안에 내재된 사회적 차별 - 빅데이터에 대한 미국 정부의 우려, 안형준, 과학기술정책 2016년(5호)

3) Executive Office of the President(2016 May), *Big Data: A Report on Algorithmic Systems, Opportunity, and Civil Rights*

- 미국 내 24,092명을 대상으로 실시한 설문조사 결과에 따르면, 응답자의 90% 이상이 빅데이터 활용에 대한 개인정보 침해에 대한 우려를 표함. 보편적인 건강보험제도가 존재하지 않는 미국의 경우에서도 빅데이터 활용에 대한 여론조사를 실시하고, 관련한 결과가 좋지 않음을 공개하고, 이를 위한 법제도 조치를 백악관에서 지시한 바 있음.
- 빅데이터 기술은 오랜 시간 축적된 방대한 사회적 데이터를 기반으로 오로지 알고리즘에 따라 분석이 수행되기 때문에 분석자의 주관이나 편견과 상관없이 객관적인 자료를 분석한다는 믿음은 틀렸다는 것을 지적함. 이러한 점을 ‘데이터 근본주의’라고 지적하고, 오류가 있음을 지적.⁴⁾ 특정 알고리즘을 입력하면 이 데이터에 반영되지 못한 이들은 소외되는 결과가 발생하고, 이러한 빅데이터가 반복되고 누적되면, 복지가 필요하지만 애초에 배제된 이들은 사실상 ‘존재하지 않는 사람들’이 되는 결과가 발생할 수도 있음. 실제와 디지털 세계 사이의 복지 사각지대가 발생하는 기술적 오류는 언제나 있을 수 있음. 이는 실제보다 데이터가 현실을 보여준다는 행정에서 기술 관료주의로 이어질 것임.
- 문제는 빅데이터에 대한 환상이 4차 산업혁명과 함께 근거도 없이 정치사회적 도그마가 되어 있는 상황에서, 이런 배제와 차별의 알고리즘에 문제를 이후 발견하고 인식하게 된다 하더라도 그것이 교정될 수 있을지 의문스럽다는 점임. 인격체가 아닌 데이터 축적기술에 지나치게 의존한 보건복지는 빅데이터 기술이 만드는 또 다른 복지 사각지대로 이어질 수 있으며 현재도 존재하는 장벽과 관료적 행정의 사각지대로 내몰린 사람들을 온정 없는 데이터셋에 가두는 결과를 가져올 수 있음.
- 이미 알고리즘이 가진 결함에 대해서는 각종 SNS에서 특정인의 글이 블로킹되거나, 특정 검색에 따라 특정인들이 걸러지거나 하는 등의 많은 예들이 있음. 하버드대학에서 나온 연구에서는 구글에 “아프리카계 미국인스러운” 이름을 넣어 검색하면 범죄자 정보를 찾아주는 회사 광고가 뜰 가능성이 높다는 사실이 밝혀졌고, 페이스북은 유색인종과 장애인의 글을 블로킹하는 알고리즘이 생성되었다는 사실이 밝혀진 바 있으며, 여성이 일자리를 검색하면 더 적은 임금의 일자리만이 더 많이 검색된다는 사실도 드러난 바 있음. 알고리즘을 만드는 주체도 인간이며,

4) 안형준(2016), '[미국] 알고리즘 안에 내재된 사회적 차별 - 빅데이터에 대한 미국 정부의 우려', 과학기술정책 2016년(5호)

인간은 편견이 있다는 사실을 잊어서는 안됨.

- 유럽 개인정보보호 감독관(EDPS)은 ‘빅데이터 문제 해결에 관한 의견서(Meeting the challenges of big data)’에서 빅데이터를 이용하려는 기관은 개인정보 처리에 어떤 원칙을 가져야 하는지를 명시하고 있음. 이 중 “투명성을 제고하기 위해서는 개인에 대해 관찰되고 추론된 개인정보가 무엇인지, 어떠한 개인정보가 처리되었는지에 대한 명확한 정보를 해당 개인에게 제공해야 하며, 개인정보의 사용 목적과 방법에 대해 보다 확실하게 고지해야 한다. 고지내용에는 목적과 방법에 대한 가정과 예상을 결정하는 알고리즘의 논리(logic)도 포함된다”고 명시하고 있음. 이러한 명시는 알고리즘에 따라 수집되고 분석된 데이터의 내용이 달라질 수 있다는 점을 감안한 것임. 이러한 유럽 개인정보보호 감독관의 의견서에 대한 복지부의 신중한 검토가 필요함.

The EU intends to maximise growth and competitiveness by exploiting big data. But the Digital Single Market cannot uncritically import the data-driven technologies and business models which have become economic mainstream in other areas of the world. Instead it needs to show leadership in developing accountable personal data processing. The internet has evolved in a way that surveillance - tracking people's behaviour - is considered as the indispensable revenue model for some of the most successful companies. This development calls for critical assessment and search for other options.

In any event, and irrespective of the business models chosen, organisations that process large volumes of personal information must comply with applicable data protection law. The European Data Protection Supervisor (EDPS) believes that responsible and sustainable development of big data must rely on **four essential elements**:

- organisations must be much more **transparent** about how they process personal data;
- afford users a higher degree of **control** over how their data is used;
- **design** user friendly data protection into their products and services; and
- become more **accountable** for what they do.

When it comes to **transparency**, individuals must be given clear information on what data is processed, including data observed or inferred about them; better informed on how and for what purposes their information is used, including the logic used in algorithms to determine assumptions and predictions about them.

<그림3-3> “Meeting the challenges of big data” the European Data Protection Supervisor (EDPS), 2015. 11

4. 효용성과 공공성

- 이 외에도 남는 문제가 많음. 심평원 사태에서 보듯이 매우 중요한 개인 정보가 한 건당 30만원의 비용으로 거래 됨. 그도 공익적 연구를 위한 목적이 아닌 기업 돈벌이에 제공. 이러한 개인정보 거래와 민영화가 보건의료 빅데이터 활성화 정책과 어떤 관련이 있을까 복지부는 해명해야 함. 사실상 복지부는 이번 심평원 사태뿐만 아니라 SK텔레콤, 약학정보원, IMS헬스 등 사태에 대한 일관성 있는 해결책을 제시하고 있지 못함. 이런 상황에서 지금의 보건의료 빅데이터 플랫폼 사업이나 추진전략 등이 누구를, 무엇을 위한 정책인지 의심스러운 것은 당연 함.
- 가정하자면 이리함. 1. 정보 제공자는 국민인데, 정보 제공자는 자신의 정보가 어떻게 이용 활용 보호 될지 알 수 없는 상황에 있다. 정부는 알 수도 없는 ‘비식별’이라는 기술적 조치를 했으니, 너의 정보가 아니다, 라고 주장한다. 2. 이 개인정보들은 돈이 되어 기업들의 돈벌이에는 매우 유용하다. 수익은 기업에게 돌아간다. 3. 이 수익을 창출하는 과정에서 내 개인정보는 ‘비식별’이라는 기술적 조치에서 예외로 ‘식별’ 될 수도 있다. 그 위험은 ‘너의 정보가 아니다’라고 했는데, 내가 지게 된다. 이건 가정이지만, 지금 정부가 추진하는 정책의 사회적 실행의 내용임. 갈등과 이해관계가 분명한 정책의 사회적 실행을 충분한 설명도 없고 동의도 없는 형태로 추진해서는 안됨.
- 정부가 혈세를 들여서 추진한다는 ‘보건의료 플랫폼 구축 사업’ 115억 원은 현행 법상 불법이지만 민간기업이 수집한 모든 데이터들까지를 연계하는 플랫폼임. 누구의 데이터이며, 누가 데이터를 보게 되는지, 데이터를 이해할 수 있는 사람들은 어디까지인지 아무런 정보가 없음. 라는 의심이 들지 않을 수 없다. 공공 기관에서 모은 개인 질병 기록과 건강 정보를 공유해달라는 주장은 지난 20년간 민간의료보험사를 비롯한 보건의료 관련 기업들의 일관된 민원상황이었음. 발제자들이 지적하고 있듯이 의료시장화의 추진 동력이 보건의료 빅데이터, 즉 개인질병/건강 정보 민영화를 축으로 추진되고 있다는 우려를 해소시킬 명확한 근거가 없으면 더욱더 이런 보건의료 빅데이터 추진은 국민적 동의를 얻을 수 없음.
- 작년에 결국 중단된 영국의 care.data.NHS 의료 정보 공유 사업에서 교훈을 얻어야 함. 국민의 동의 절차를 무시했다는 점. 비식별화 기술적 조치가 해결하지 못한 익명화 논란이 결국 100만 명의 탈퇴를 야기했고, 이는 보수당 정권 자체가

흔들리고 큰 사회적 혼란과 비용을 초래했다는 점. 우리가 얻을 교훈은 국민 동의 없이 추진되는 이번 보건의료 빅데이터 사업이 결국 건강보험에 대한 신뢰에 기반해 제공한 개인정보 수집 권한을 남용하는 정부 당국에 대한 불신으로 이어질 것이라는 점. 익명화가 아닌 비식별 조치를 내세우고 있는 정부의 안일한 기술행정 관료주의가 마침내 한 개인에게 강요할 희생과 삶의 파탄. 그리고 그러한 파탄은 보다 어려운 사회 계층에게 집중 될 것이라는 점임.

- 동의 절차 없는 갑질정부, 묻지마 행정 정권은 박근혜로 충분. 복지부는 지난 부패정권의 추진 전략을 베끼고 열기설기 끼워 맞추는 관료행정을 중단하고, 기업의 먹거리 전략에 밀밥 대주는 정책을 사회적으로 실행해서는 안됨. 정책은 사회적 실행을 위한 것이며, 이러한 사회적 실행의 결과가 어떤 결과를 낳을지에 대한 다양한 연구와 토론이 더 많이 필요함. 복지부는 일단 박근혜 정권에서 추진된 부문별한 행정 가이드라인이 현행법과 불일치하고 있는 문제를 정리해야함. 이는 국회에서 지금 해야 할 일이기도 함.
- 마지막으로 저성장의 돌파구가 될 것이라고 예견하고 있는 ‘4차 산업혁명’에 대한 거품을 거둬내는 일이 필요함. 문재인 정부도 나서서 선전하고 있는 4차 산업혁명의 실제 내용은 박근혜 정부 시절 규제완화를 위한 레토릭과 다르지 않음. 4차 산업혁명에서 말하는 산업의 실체는 고위험 고부가가치 사업임. 이런 빠른 자본의 투자수익률을 위한 산업 패러다임을 보건복지에 가져와 사용하는 것은 문제임. 고부가가치 창출은 고 위험을 전제하고 있으며, 이는 규제완화를 조건화하고 있다는 점을 간과해서는 안됨. 보건의료 분야에서 빅데이터가 4차 산업혁명에서 말하는 ‘혁신적 기술’일지는 몰라도 “환자의 사전 동의 없이 공유되는 데이터의 양”이 많아질수록 가능한 ‘혁신 기술’ 일 경우, 사회는 이를 선택할 것인가? 거부할 것인가 논의해야 하며 이는 사회 구성원 모두의 권리이기도 함. 국회는 이러한 점에서 민의를 대변하는 기구로서 보건의료 빅데이터 추진에 대한 지협적이고 행정 기술적인 논의가 아닌, 보다 민주적인 방식과 절차를 모색하는 대의기관으로의 역할을 해야 할 필요가 있음.

토론문

이찬진 / 참여연대 사회복지위원회 실행위원, 변호사

1. 민감정보인 건강정보의 처리 실태

- 건강보험제도 하의 의료기관을 중심으로 한 처방정보 전부와 진료정보(그 중 일부 범위에 관하여는 다툼의 여지가 있어 보임)에 관한 사항들은 개인정보보호법 제23조에서 정한 ‘민감정보’로 당사자의 동의를 받았거나 ‘법령에서 민감정보의 처리를 요구하거나 허용하는 경우’에만 예외적으로 정보처리자에게 민감정보 처리를 허용하고 있음. 현재 건강보험 운용 실태 상으로 환자의 건강정보는 급여청구를 위하여 환자가 이용한 1)의료기관 및 조제약국 — 2)심평원/-- 건보공단으로 전송되고 있음.
- 1)에서 2)로 정보가 이동하는 과정에서의 수탁처리업체의 민감정보에 관한 불법 처리와 상업적 이용이 문제된 사례가 약학정보원 및 병원에 급여심사 및 청구 관리 프로그램을 납품하고 수탁처리를 하는 수탁처리업체가 불법으로 민감정보를 보유하고, 이것을 ‘비식별’처리의 외형을 만들어 미국에 본사를 둔 외부 영리사업자에게 유상으로 처분하였고, 동 정보들은 미국 본사를 통하여 전 세계에 유상으로 제공되고 있는 것이 문제가 되었음. 이와 관련하여 관련 업체 및 임직원들이 기소되어 현재 형사 재판이 진행 중인 사례가 있고, 한편으로는 이와 같은 불법 정보처리를 한 약학정보원과 00 수탁처리업체의 불법행위에 대하여 손해배상 청구소송을 제기하여 최근 1심 판결이 선고된 사례가 있음(이하‘IMS 사건’이라 함).

- 더 심각한 것은 2)의 단계에서 급여심사 및 급여 지급을 위하여 전송받은 환자의 건강정보를 집적하고 있는 ‘심평원’이 2014년 7월부터 올해 8월까지 비식별화된 정보를 지속적으로 제공한 것임. 민간보험사 등에 여러 차례에 걸쳐서 6,420만 명의 건강정보 등이 포함된 데이터를 유상으로 제공했다는 것임. 위 두 가지 유형을 기준으로 하여 법률적 문제점과 개선방안을 발제문을 중심으로 검토하고자 함.

2. ‘IMS 사건’ 관련 검토

- 위 사건의 핵심적인 문제점은 우리나라 개인정보보호법상 허용되지 않은 민감 정보인 건강보험 관련 처방정보와 진료정보가 의료인, 약국에서 심평원으로 전송 되기 전에 일종의 수탁처리자들에게 전산 프로그램을 통하여 집적,관리되고 있다는 점임. 이 유통과정은 개인정보보호법이나 국민건강보험법, 의료법, 약사법 등 관련 법 그 어디에도 허용된 바 없는 것임.
- 따라서, 민사 소송 판결에서 재판부는 개인정보처리자가 ‘비식별화’만 하면 그 정보가 개인정보가 되지 않는 것으로 판단하고 있는데 가장 중요한 쟁점인 약국이 처방정보와 진료정보를 제3자에게 처분할 수 있느냐에 관하여는 오히려 수탁처리업자가 약국의 동의만 받으면 되는 것으로 판단하는 오류를 보이고 있으며 이는 항소심에서 주요한 쟁점이 될 것으로 예상됨. 나아가 수탁처리자가 불법으로 보유하고 있는 개인정보를 통상적인 수준의 비식별화만 하면 그 자체로 개인정보의 법적 지위를 상실하므로 이를 제3자에게 유상으로 처분하여도 불법행위가 아니라는 식의 판단은 더욱 심각한 문제가 될 수 있음.
- 관련 형사 사건의 공소 제기로 현재 형사 재판 중인 사건의 경우에도 가장 중요한 쟁점인 약국에서 약학정보원에게 정보제공동의를 하였다는 점 자체가 개인정보보호법 위반이 되는데 이 부분이 정면으로 다투지지 않고 있다는 점에 있음 즉, 1차적인 민감정보 대량 유통의 길을 사실상 묵인하는 결과가 되고 있는 것이 현실임.

3. 심평원 빅데이터 정보 제공 건

- 가장 중요한 것은 ‘심평원’이 급여심사, 평가를 하면서 보유하게 된 환자들의 개인정보를 보유·집적·조합 등 가공할 수 있는 법적 근거가 있는지를 살펴보아야 함. 적어도 국민건강보험법상 ‘심평원’이 이와 같은 권한을 부여받은 바는 없음.
- 공공데이터의 제공 및 이용 활성화에 관한 법률 제1조는 “공공기관이 보유·관리하는 데이터의 제공 및 그 이용 활성화에 관한 사항을 규정함으로써 국민의 공공데이터에 대한 이용권을 보장하고, 공공데이터의 민간 활용을 통한 삶의 질 향상과 국민경제 발전에 이바지함을 목적으로 한다”⁵⁾라고 규정하고 있는데 정작 중요한 것은 심평원이 급여심사·평가를 위하여 제공받은 건강정보를 심평원이 ‘보유’ 및 ‘처리’할 수 있는 법률적인 근거는 전혀 없다는 점임. 따라서 심평원이 갖고 있는 개인 건강정보를 비식별화 할 수 있는 법률상의 권한은 현재 없는 것임. 이와 같은 개인정보 처리 차원의 비식별화 조치를 하려면 환자들의 개별적인 동의(정보 보유 및 처리 모두에 대하여)를 받아야 함. 이 점에 있어서 발제자들의 의견에 모두 동의함.
- ‘심평원’은 급여심사 및 평가가 완료되면 당연히 동 정보는 개인정보보호법에 따라 폐기되어야 하며, 만일 법령상의 근거가 신설되어 심평원이 개인정보처리자의 지위를 확보한다고 하여도 적어도 개인정보보호법에서 정한 정보주체의 권리를 보장하기 위한 제반 규정을 준수하여야 할 것임. 따라서 ‘심평원’이 법률상 권한도 없는 상태로 보유하고 있던 개인건강정보를 ‘비식별화’한 것도 개인정보보호법 위반으로 판단되고, 이를 유상으로 민간 보험사 등에게 제공한 것도 정보 주체의 동의를 받지 아니한 것으로서 모두 법 위반으로 판단됨.

5) 관련 법률로 국가정보화기본법이 있으며, 동법 제3조 2. "정보화"란 정보를 생산·유통 또는 활용하여 사회 각 분야의 활동을 가능하게 하거나 그러한 활동의 효율화를 도모하는 것을 말한다.

3. "국가정보화"란 국가기관, 지방자치단체 및 공공기관이 정보화를 추진하거나 사회 각 분야의 활동이 효율적으로 수행될 수 있도록 정보화를 통하여 지원하는 것을 말한다. 라고 규정하고 있음.

4. 보건복지부 건강보험 빅데이터 관련 ‘개인정보 비식별 조치 가이드라인’의 문제점

- 위에서 본 바와 같이, ‘심평원’이나 ‘건보공단’이 보유한 환자들의 개인정보는 정보주체인 환자의 동의를 받지 않은 것이므로 급여심사 평가 및 급여지급 목적이 완료되면 폐기되어야 하는 것이고(만일 이를 보유하려면 국민건강보험법이나 개인정보보호법에 관련 근거 규정을 신설하여야 함), 현재 보유한 데이터를 비식별화하는 것 자체가 정보주체의 동의를 받지 않는 것이어서 법률 위반임. 따라서 이와 같은 현행 법률을 무시한 채 보건복지부가 정한 ‘비식별화 가이드라인’은 위법한 공권력 행사로서 국가 불법행위를 구성한다고 해석될 가능성이 매우 높음.

5. 개선 방향

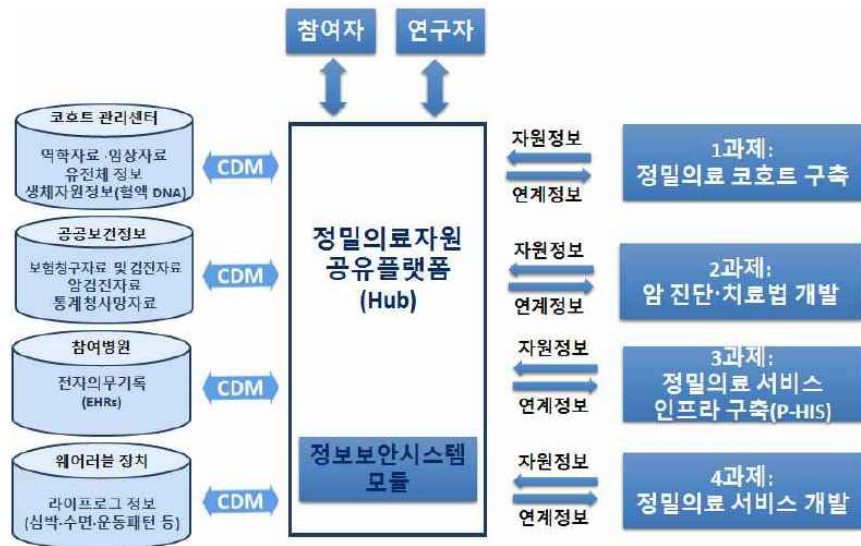
- 발제자의 용에 대해 기본적으로 동의하면서 다음 사항을 보충하고자 함.
- 먼저, 의료기관/약국에서 심평원으로 가는 급여청구 관련 정보의 처리와 관련하여 수탁처리 프로그램 및 수탁처리자는 원칙적으로 제3의 공공기관이 담당하도록 법률 개정이 필요함. 현재와 같은 약국 및 의료기관 급여/청구 관리 프로그램이 필요한 것이 현실임을 인정한다고 하여도 이들이 처리하는 정보는 환자들의 민감정보라는 점에서 그 정보 처리자는 반드시 공공이 되어야 할 것이며, 이 단계에서 그 어떠한 형태의 정보처리 및 영리사업화는 금지되어야 할 것임.
- 심평원이나 공단이 업무상 취급하는 처방정보, 진료정보에 관하여 심평원과 공단의 정보처리권한과 보존기간을 분명히하고, 비식별화 처리 주체 역시 공공이 공적으로 관리·운용하며, 그 이용 목적 역시 공공적인 성격의 범위 내에서 운용될 필요가 있을 것임. 이와 같이 공적으로 가공되어 공표된 빅데이터를 기초로 하여 필요한 경우 영리 민간 사업자들이 2차적으로 이를 활용하는 것은 가능할 것임. 현재와 같은 방임형 빅데이터 산업화는 정보주체의 인권은 심각하게 침해할 수 있음.

토론문

김병수 / 성공회대학교 열림교양대학 교수

1. <유전체 · Health-ICT 융합 기반 정밀의료 기술개발사업> 중 일부.

- 현재 논란이 되고 있는 의료 빅데이터 사업은 지난 정부가 급조한 정밀의료 사업의 세부 사업으로 진행되고 있음. 약간의 변동이 있겠지만 현 정부에서도 아래와 같은 형태로 추진될 가능성이 높음.
- 개인의 의료기록에서부터 라이프 스타일, 위치정보, 거주 환경 심지어는 DNA 정보까지 수집 통합 관리하려는 시도로 보임. 국가가 개인의 거의 모든 정보를 수집하는 계획을 수립하면서 사회적 논의는 전혀 없었음.
- 목적 : 한국인 정밀의료 인프라 구축 · 활용을 통한 맞춤형 의료·건강관리 서비스 확대 (정부안) (5,063.4억 원 - 국고 3,992.6억 원, 민간 1,070.8억 원; 5년)
 - 세부사업 1 : 한국인 정밀의료 코호트(KPM) 구축 및 개방(1,370.8 억 원)
 - 세부사업 2 : 정밀의료 기반 진행성 암 진단 · 치료기술 개발(1,383.5 억 원)
 - 세부사업 3 : 의료 빅데이터 통합 · 분석 시스템(P-HIS) 구축(849.9억 원)
 - 세부사업 4 : 정밀의료 기반 건강관리서비스 개발(533.7억 원),
AI 기반 진단 · 치료 지원 솔루션 개발(848.7억 원)



<표4-1> 수집 대상 정보 및 내용

	수집 대상 정보 및 내용	대상 인원
세부사업 1	(기본정보) 참여 동의서, 인구학적 특성, 연락처 등 (역학 및 임상정보) 질병과거력, 치료현황, 약물력, 가족력, 여성력, 사회사리요인, 건강상태 정보, 자가보고 측정정보(흡연, 음주, 식생활, 신체활동력 등), 신체계측, 혈액 및 소변검사 정보, 임상검사 정보(생체시료) 혈장, 혈청, genomic DNA, 소변, 모발, 생체조직 등(유전정보) 염기서열정보(NGS), 유전체 칩 등 사용(Life-Log Data) 센서 기반 행동 및 라이프스타일 관찰정보(수면, 식이, 신체활동, 생활습관정보, 위치정보, 활동정보, 심박수 등)	10만 명 (최대 50만 명)
세부사업 2	생검·DNA 추출 후 300여 개 유전자 검사 시판 중 또는 시판 예정인 표적치료제, 임상시험 대기 중인 후보물질 등 최대 20종 약물에 대한 임상시험 수행 이차연구에 동의한 환자의 시료를 별도로 저장하고 관리	암환자 1만 명 유전체 DB 2천 명 임상시험 정보 DB

○ 2016년 KISTEP 예비타당성 분석 후 752.3억 원(630.8억 원)으로 제안.

2. 기존 사업과의 관계

<표4-2> 기존 사업과의 관계

부처	사업명	사업 기간
보건복지부(4개)	임상연구인프라조성사업	'08~계속
	질환극복기술개발사업	'13~계속
	첨단의료기술개발사업	'08~'20
	형질분석연구사업	'12~계속
질병관리본부(4개)	질병유전체정보화사업	'12~계속
	한국인유전체사업	'12~계속
	한국인유전체역학조사사업	'01~계속
	한국인체자원은행사업	'08~계속
미래창조과학부(3개)	바이오·의료기술개발사업	'04~계속
	빅데이터 활용 스마트서비스 시범사업	'14~계속
	중증질환자 after-care 기술개발 및 실증사업	'15~'17
산업통상자원부(2개)	산업핵심기술개발사업	'09~계속
	창의산업미래성장동력사업	'12~'18
다부처(2개)	포스트게놈다부처유전체사업	'14~'21
	질환유전자분석플랫폼기술개발사업	'15

- 이미 2001년부터 여러 연구 사업에서 개인의 유전체를 분석 보관 분양하고 있음. 2008년부터 시작한 Biobank사업에서만 최소 67만 명의 검체를 확보한 것으로 보임(세계최대). 여기에 추가로 수십만 명의 DNA를 수집 분석해야하는 근거를 밝혀야 함. 따라서 기존 사업에 대한 면밀한 평가 후 추가 사업 여부에 대해 논의해야 함. 이러한 평가도 없이 추가로 10~50만 명의 DNA 수집하려는 계획은 타당성이 없음.
- HGP 이후 postgenomics라는 이름 아래 다양한 연구들이 진행되고 있으나 의료적 유용성 등을 둘러싼 논란은 지속되고 있음. 정밀의료에 대한 의료적 혜택은 과장돼 있는 반면 개인의 유전체 프라이버시(genomic privacy)는 위협받고 있음. 의료적 성과는 불확실⁶⁾한 반면 일부 기업들을 위한 시장 창출(NGS분석 등)은 구체적임.

3. <비식별화 가이드라인>, <포괄적 동의서>, 홍보로 해결?

- 개인의 모든 정보를 수집해 분석 저장 활용하는 것도 문제지만 법적인 기반 없이

6) □ 보건복지부는 정밀의료 기술개발을 통해, '25년 기준으로 다음과 같은 기대효과가 있을 것으로 예측 ○ 건강수명 3년 연장(73세 → 76세) 및 현재의 절반 수준으로 의료비* 증가율 감소(7% → 3%)

○ 3대 전이암(폐암, 위암, 대장암) 5년 생존율 6% 증가(8.4% → 14.4%) (2016.8.10. 복지부 보도자료)

가이드라인이나 informed consent, IRB를 통해 다양한 정보를 손쉽게 수집하려는 시도는 더욱 문제임(영국의 UK biobank는 2000년부터 시작했지만 시민참여, 전문가 자문, 윤리적 쟁점 연구, 거버넌스 시스템 구축 등으로 실제 검체는 2006년부터 수집했음 --> 50만 명).

- 개인의 DNA 변이, 환경, 라이프 스타일을 고려한 질병의 치료와 예방을 목적으로 하는 정밀의료는 ‘개인의 DNA 변이 정보’ 확보가 필수적임. 따라서 예산이 확보되면 DNA 수집을 추진할 것으로 보임.
- 정밀의료 사업은 DNA 변이* 환경을 고려한 단순한 맞춤형 의료 시스템이 아니라 우리나라의 보건의료 시스템 전체에 영향을 주는 사업으로 추진 ‘절차’에 대한 논의가 아닌 ‘사업 자체’에 대한 사회적 논의와 사회적 합의가 필요함.

토론 4

토론문

김록영 / 건강보험심사평가원 빅데이터부 부연구원

토론 5

토론문

오상윤 / 보건복지부 의료정보정책과 과장

토론 6

토론문

정영수 / 행정안전부 개인정보보호정책과 사무관

토론회 심평원 및 약학정보원 개인질병정보 판매 행위로 본
현 정부의 보건의료 빅데이터 추진 전략의 문제점

발행일 2017. 11. 27

발행처 김상희 의원, 남인순 의원, 정춘숙 의원, 윤소하 의원, 참여연대, 건강과대안, 인도
주의실천의사협의회, 진보네트워크센터, 무상의료운동본부, 경제정의실천시민연합,
건강세상네트워크