

토픽모델링과 에고 네트워크 분석을 활용한 스마트 헬스케어 연구동향 분석

윤지은¹ · 서창진^{2*}¹한양대학교 일반대학원 경영학과²한양대학교 경영학부

Research Trend Analysis on Smart healthcare by using Topic Modeling and Ego Network Analysis

Jee-Eun Yoon¹ · Chang-Jin Suh^{2*}¹Dept. of Business Administration, Graduate School, Hanyang University, Seoul 04763, Korea²School of Business, Hanyang University, Seoul 04763, Korea

[요 약]

스마트 헬스케어는 ICT 분야와 의료서비스 분야가 융·복합된 분야로 다양한 분야에서 학제 간 융·복합 연구가 활발히 이루어지고 있다. 본 연구는 토픽모델링(Topic Modeling)과 에고 네트워크 분석(Ego Network Analysis)을 활용하여 스마트 헬스케어 연구동향을 살펴보는 데 그 목적이 있다. 이를 위해 2001년부터 2018년 4월까지 Scopus에 게재된 2,690편을 대상으로 텍스트 분석, 각 기간별 빈도분석, 토픽모델링, 워드 클라우드, 에고 네트워크 분석을 수행하였다. 토픽 모델링 분석 결과 8개의 주요 연구토픽이 도출되었다. 8개 주요 연구토픽은 “AI in healthcare”, “Smart hospital”, “Healthcare platform”, “blockchain in healthcare”, “Smart health data”, “Mobile healthcare”, “Wellness care”, “Cognitive healthcare” 순으로 나타났다. 토픽모델링 결과를 보다 심도 있게 살펴 보기 위해 연구토픽별 에고 네트워크 분석을 하였다. 이를 통해 스마트 헬스케어 연구동향을 파악하고, 향후 연구의 방향성을 수립하는데 시사점을 제시하고자 한다.

[Abstract]

Smart healthcare is convergence of ICT and healthcare services, and interdisciplinary research has been actively conducted in various fields. The objective of this study is to investigate trends of smart healthcare research using topic modeling and ego network analysis. Text analysis, frequency analysis, topic modeling, word cloud, and ego network analysis were conducted for the abstracts of 2,690 articles in Scopus from 2001 to April 2018. Topic Modeling analysis resulted in eight topics, Topics included “AI in healthcare”, “Smart hospital”, “Healthcare platform”, “Blockchain in healthcare”, “Smart health data”, “Mobile healthcare”, “Wellness care”, “Cognitive healthcare”. In order to examine the topic modeling results core deeply, we analyzed word cloud and ego network analysis for eight topics. This study aims to identify trends in smart healthcare research and suggest implications for establishing future research direction.

색인어 : 스마트 헬스케어, 텍스트 분석, 토픽모델링, 워드 클라우드, 에고 네트워크 분석

Key word : Smart Healthcare, Text Analysis, Topic Modeling, Word cloud, Ego Network Analysis

<http://dx.doi.org/10.9728/dcs.2018.19.5.981>



This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0/>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

Received 15 May 2018; Revised 25 May 2018

Accepted 25 May 2018

*Corresponding Author; Chang-Jin Suh

Tel:

E-mail:

I. 서론

2016년 1월 다보스포럼(WEF, World Economic Forum)에서 클라우드 슈밥(Klaus Schwab) 교수가 4차 산업혁명이라는 화두를 던지면서[1] 전 세계적으로 이를 기반으로 신성장동력을 찾아 지속 가능한 경제성장을 이어가려는 움직임이 활발해지고 있다[2]. 4차 산업 혁명은 정보통신기술(ICT; Information and Communication Technology)을 필두로 다양한 산업 간 융·복합 기술 혁신들이 이루어지고 상호작용하면서 빠르게 진화해 가는 인류 발전의 새로운 장으로 인식된다. 이는 엄청난 속도와 과급력을 가지고 지금까지 상상하지 못했던 새로운 가치 창출의 생태계를 만들어 낼 것이다[3]. 이와 같은 현상을 MIT의 에릭 브린올프슨(Eric Brynjolfsson) 교수와 앤드루 맥아피(Andrew McAfee) 교수는 ‘제 2의 기계시대(The second machine age)’라고 일컬으며, 자동화로 인한 디지털 기술의 영향력이 ‘완벽한 힘(Full force)’을 갖추고 ‘기존에 존재하지 않았던 새로운 것(unprecedented things)’들을 만들어내는 변곡점(inflexion point)의 시기에 있다고 하였다[1].

4차 산업혁명의 중심에 있는 ‘스마트 헬스케어’분야 역시 4P(Preventive, Predictive, Personalized, Participatory)로 대변되는 헬스케어산업 전반의 패러다임 변화를 이끌어 내는 촉매제 역할을 할 것이다[4]. 글로벌 조사 전문기관인 Statista에 의하면 전 세계 스마트 헬스케어 시장규모는 2013년 608억 달러에서 2017년에는 1,359억 달러, 2020년에는 2,333억 달러로 연평균 22% 성장할 것으로 전망하고 있다[5]. 미국, 유럽, 일본 등 선진국들은 전략적으로 ICT와 다양한 헬스케어산업의 융·복합 기술들을 적극적으로 개발 육성하여 고부가가치 헬스케어 신산업 분야를 선점하려는 움직임이 활발히 나타나고 있다[2]. ‘스마트 헬스케어’는 병원 및 의료서비스산업, 정보통신업, 제약, 의료기기 등 연관 제조업 등이 밀접하게 연계된 최첨단 융·복합산업 분야로 전반적인 국가경쟁력 우위를 확보하기 위해서도 적극적인 투자확대 및 적절한 투자전략 수립이 필요한 분야이다[6].

본 연구에서는 ‘스마트 헬스케어’분야의 최근 주요 중점 연구동향을 파악하고 시사점을 도출하기 위해 Scopus에서 2001년부터 2018년 4월까지 게재된 영문 연구 초록을 분석대상으로 텍스트 분석, 빈도분석, 토픽모델링, 워드 클라우드, 에고 네트워크 분석을 수행하였다.

II. 이론적 배경

2-1 스마트 헬스케어

스마트 헬스케어는 상황인식(Context-aware) 네트워크와 스마트 도시(Smart city)의 ICT 인프라 감지(sensing infrastructure)와 그 기술들 그리고 모바일 헬스(m-health)와 원격의료(telemedicine)의 개념이 혼합되어 유비쿼터스(Ubiquitous)개념보다 더 확장된 개념으로 새롭게 정의되고 있다[7]. 스마트 헬스케어는 사물인터넷(IoT; Internet of Things), 빅데이터(Big

data), 클라우드 컴퓨팅(Cloud computing), 인공지능(AI; Artificial Intelligence) 등과 헬스케어 관련 산업이 융·복합한 분야[8]로서 개인건강관리, 의료정보, 의료기기 및 다양한 형태의 시스템과 플랫폼 등을 아우르는 분야이다. 스마트 헬스케어의 목표는 예측 및 맞춤형 의료와 건강관리서비스를 제공하는 것으로 개인이 소유한 스마트 폰, 웨어러블 기기, 클라우드, 병원 정보 시스템 등에서 확보된 생체신호 정보, 생활 습관, 생활환경 정보, 건강검진, 의료이용 정보, 유전체 정보 등 방대하고 다양한 데이터 수집과 분석을 통해 제공된다[9].

스마트 헬스케어 산업의 연구동향을 파악하기 위한 연구로는 데이터 분석기술(헬스케어 플랫폼, 인공지능 등)을 기반으로 한 헬스 빅데이터 사례연구[10], 환자중심 서비스 제공을 위한 스마트 시스템과 이와 관련된 정보시스템 연구현황 [11], 스마트 헬스케어 범위를 좁혀 심전도 측정 기술 및 기기의 개발 동향 국내·외 사례연구[12], 스마트의료 환경에서 보안위협과 이에 대응하는 보안 기술에 대한 국내·외 연구동향과 표준화 동향[13] 등이 있었다. 하지만 현재 활발히 수행되고 있는 스마트 헬스케어 분야 전반의 주요 중점 연구주제를 파악하거나 각 연구 주제별 최신 연구동향의 변화 양상을 종합적으로 분석한 연구는 수행된 바 없는 것으로 파악된다.

2-2 토픽 모델링(Topic Modeling)

토픽모델링은 비정형의 방대한 문서(original texts)의 단어들을 분석해서 문서 안에 숨어 있는 주제(themes)들을 찾아내고, 각 주제들이 서로 어떻게 연결되었는지, 시간이 흐름에 따라 어떻게 변화하는지를 분석하는 통계기반 방법론[14]으로 이슈를 도출하기 위한 연구에서 널리 활용되고 있다. 토픽모델링에서 가장 많이 사용하고 있는 생성 확률모델(generative probabilistic model)인 잠재 디리클레 할당(LDA; Latent Dirichlet Allocation) 모델[15]은 텍스트 자료의 의미구조를 파악하기에 적합한 수학적이고 통계적인 방법이다[16]. 즉, 잠재적인 토픽(latent topic)들이 무작위로 혼합(random mixtures)되어 문서를 이루고 있다고 가정하고, 각 토픽의 특성은 문서상에 분포해 있는 관측 가능한 단어들의 패턴을 통해 추론된다[15]. 토픽 모델링은 호텔·관광·문화정보학·인적자원개발(HRD) 등 다양한 분야의 문헌, 소셜 미디어(트위터, 네이버 블로그 등), 뉴스 데이터 등 방대한 텍스트에서 정보를 얻고자 하는 연구에서 다양하게 활용되고 있다. 본 연구는 LDA 기반 토픽모델링을 통해 스마트 헬스케어 분야 연구동향을 파악하고자 한다.

2-3. 에고 네트워크 분석(Ego Network Analysis)

에고 네트워크 분석은 텍스트 네트워크 분석(Text Network Analysis)의 한 방법이다. 텍스트 네트워크 분석은 텍스트를 분석하는 다양한 방법 중 하나로 언어와 언어 간 관계를 분석한다는 측면에서 언어 네트워크 분석이라고도 할 수 있다. 텍스트 네트워크 분석은 텍스트에 직접적으로 나타나 있는 주요개념들 간 연결 패턴을 분석을 하는 것이고, 언어의 의미 네트워크 분석은 텍스트에서 숨어 있는 맥락을 파악해서 전달하고자 하는

의미를 분석하는 것이다[17]. 텍스트 네트워크 분석의 구성요소는 노드(node)와 링크(link)가 있다. 노드는 상호 고유한 속성을 가지는 키워드와 사람 등을 의미하고, 링크는 노드들 간 연결관계를 나타낸다.

텍스트 네트워크 분석은 목적에 따라 에고 네트워크 분석을 할 수 있다. 에고 네트워크 분석은 네트워크 내 특정 노드를 에고(ego)¹⁾로 설정하여 에고와 직접 링크(연결 관계)를 형성한 이웃 노드(alter)²⁾들을 대상으로 분석하는 것이다[18]. 네트워크 분석은 미래예측이나 트렌드를 반영하기 어려운 기존의 정성적인 분석을 보완하는 하나의 방법으로 제시되고 있다[19]. 본 연구에서는 스마트 헬스케어 토픽모델링 결과를 보다 효과적으로 분석하기 위해 각 토픽별 할당확률이 높은 노드 중심으로 에고 네트워크 분석을 시행하였다.

III. 연구 방법

3-1 분석대상

본 연구는 ‘스마트 헬스케어’ 분야의 연구동향을 파악하기 위해 Elsevier사의 Scopus에서 2001년부터 2018년 4월까지 약 18년간 게재된 2,878편의 논문(Article, Conference Proceedings) 데이터를 수집하였다. 텍스트 분석을 위해 영문(English)으로 작성된 논문들을 대상으로 하고, 저자정보가 없는 논문 157편, 초록이 없는 논문 22편과 중복된 논문 9편을 제외하여 총 2,690편을 분석 대상으로 선정하였다.

3-2 분석방법

본 연구는 <그림 1>과 같이 Scopus에서 수집한 논문의 영문 초록을 대상으로 텍스트분석, 빈도분석, 토픽모델링, 워드 클라우드, 에고 네트워크 분석을 하는 순서로 연구를 진행하였다. 즉, ‘스마트 헬스케어’ 관련 논문 초록을 텍스트 분석하여 기간별 핵심 단어의 빈도를 살펴보고, LDA기반의 토픽 모델링으로 이 분야에서 수행되고 있는 주요 중점 연구주제(토픽)들을 도출하고, 각 연구주제별 워드 클라우드와 에고 네트워크 분석을 통해 ‘스마트 헬스케어’ 분야 최신 연구동향들을 파악하고자 한다.

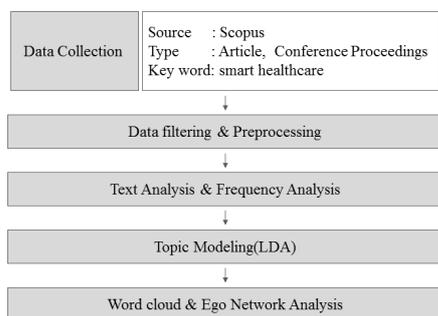


그림 1. 연구 분석 흐름도

Fig. 1. Analysis Flowchart of Study

- 1) 개인, 집단, 조직, 키워드 등 분석하고자 하는 1개 혹은 다수의 특정 노드
- 2) 에고 노드와 연결 관계가 '1'인 인접노드(adjacent node)

데이터 전처리 단계에서는 <표 1>과 같이 소문자 통일, 약어와 단·복수형, 띄어쓰기 등을 표준화하고, 특수문자와 stop words 등 불필요한 속성을 제거하는 필터링 작업을 하여 말뭉치(corpus)를 작성하였다. 사용자 정의 사전(복합명사, 유의어)과 불용어사전은 2,690편 논문 키워드 기반으로 생성하였다. 분석에는 NetMiner 4.3을 사용하였다. 논문 초록의 텍스트를 분석하여 단어들을 추출하고, 빈도분석, 토픽모델링, 워드 클라우드, 에고 네트워크 분석을 시행하였다. 빈도분석은 기간별 연구동향을 파악하기 위해 <표 2>과 같이 2001년 기준으로 5년 간격으로 2001년-2005년, 2006년-2010년, 2011년-2015년으로 구분하고 마지막으로 2016년-2018년 4월 까지 구분에서는 2016년 1월 WEF에서 ‘4차 산업혁명’이 언급된 이후의 연구동향을 파악하고자 하였다. LDA 기반으로 토픽모델링을 하여 토픽을 추출하고, 각 토픽별 워드 클라우드와 에고 네트워크 분석을 하여 주요 토픽별로 구체적인 연구내용들을 보다 효과적으로 분석하고자 하였다.

표 1. 데이터 전처리 및 사용자 정의 사전

Table. 1. data preprocessing and custom dictionary

data preprocessing	lower case	Smart Healthcare → smart healthcare
	singular	devices → device
	abbreviated form	iot → internet of technology
custom dictionary	synonym	medical doctor → doctor, surgeon, physician, medical professional etc.
	stop words	method, study, ieee etc.

IV. 연구결과

4-1 빈도분석

빈도분석에서 도출된 단어는 기간별 연구된 논문 중 얼마나 많은 논문에서 등장했는지를 나타내며, 빈도가 높은 단어일수록 해당 기간 중 중점적으로 연구된 주제일 가능성이 높다. 본 연구에서는 수집한 논문 2,690편을 <표 2>와 같이 2001년-2005년, 2006년-2010년, 2011년-2015년, 2016년-2018년 4월로 기간을 나누어 각 기간별 어떤 주요 연구주제들이 등장했는지 빈도분석을 하였다. 각 기간별 고빈도 상위 30개 주제를 뽑아 새롭게 등장한 주제는 음영처리하고, 순위권에 벗어난 주제는 대각선을 그었으며, 생성과 소멸을 반복한 연구주제는 음영과 대각선을 모두 표기하였다.

표 2. Scopus 추출 논문 수

Table 2. Article number of Scopus

year	document	document rate
2001-2005	74	3%
2006-2010	367	14%
2011-2015	1142	42%
2016-2018	1107	41%

1) 빈도분석 결과 (2001-2005)

2000년대 초인 2001년부터 2005년 발표논문에서는 smart card, medical prescription, cryptography, medical doctor, security 와 같은 주제어들의 빈도가 높아 이와 관련된 연구들이 중점적으로 수행되었다고 볼 수 있다. prevention, risk assessment, smart material, wearable technology, public key infrastructure(PKI), rehabilitation, nano technology(NT), convergence는 다음 시기에 순위권 밖으로 밀려난 연구주제들이다.

표 3. 빈도 분석 결과(2001-2005)

Table 3. Result of Frequency Analysis(2001-2005)

Word	Freq	Word	Freq
smart card	49	human computer interaction(HCI)	15
medical prescription	48	pharmacy	15
cryptography	47	machine learning	15
medical doctor	37	context awareness	14
security	32	emergency	14
telemedicine	29	remote network monitorin	14
smart home	27	cost efficiency	13
policy	22	information and communication technology(ICT)	13
health data	22	personal healthcare system(PHS)	13
prevention	21	wearable technology	12
material	21	public key infrastructure(PKI)	12
chronic condition	19	rehabilitation	11
risk assessment	17	nano technology(NT)	11
data integration	16	sensor technology	11
smart material	16	convergence	10

2) 빈도분석 결과 (2006-2010)

2006년부터 2010년까지의 연구에서는 medical doctor, cryptography, security, medical prescription이 여전히 주를 이루었고, smart phone, wireless network, platform, algorithm, functionality, hear rate가 새로운 연구주제로 급상승했다. lifestyle disease, bio technology는 새로운 연구주제로 등장했다가 다음시기에 순위권 밖으로 밀려났고, material, smart card,

telemedicine, pharmacy, information and communication technology(ICT), health data 또한 다음 시기에 순위권 밖으로 밀려났다. 이를 통해 2001년부터 2010년까지의 기간별 주요 연구주제에 상당한 변화가 있었던 것을 알 수 있다.

표 4. 빈도 분석 결과(2006-2010)

Table 4. Result of Frequency Analysis(2006-2010)

Word	Freq	Word	Freq
medical doctor	132	sensor technology	62
cryptography	126	platform	60
security	108	data integration	59
smart phone	102	lifestyle disease	57
medical prescription	99	bio teehnology	56
wireless network	98	smart card	56
emergency	86	telemedicine	55
smart home	86	machine learning	54
remote network monitoring	78	pharmacy	53
cost efficiency	73	algorithm	51
context awareness	72	information and communication technology(ICT)	50
chronic condition	67	policy	50
personal healthcare system	67	functionality	49
human computer interaction(HCI)	66	health data	48
material	65	heart rate	47

3) 빈도분석 결과 (2011-2015)

2011년부터 2015년 연구에서는 smart phone이 가장 중점적으로 연구되었고, internet of things(IoT), real time data, feature extraction, m-health, cloud computing, energy, emotion preception, privacy가 새로운 주요 연구주제로 등장했다. 이를 통해 이 기간 헬스케어 분야와 연관분야 간의 첨단기술 융·복합 관련 연구주제들이 새롭게 부상하고 있음을 확인할 수 있다. risk assessment는 새롭게 등장했다가 순위권 밖으로 밀려났으며, data integration, sensor technology, functionality 또한 순위권 밖으로 밀려났다.

표 5. 빈도 분석 결과(2011-2015)

Table 5. Result of Frequency Analysis(2011-2015)

Word	Freq	Word	Freq
smart phone	579	data integration	201
cryptography	429	real time data	199
medical doctor	382	feature extraction	198
medical prescription	348	m-health	188
context awareness	306	chronic condition	187
wireless network	301	personal healthcare system	180
machine learning	267	algorithm	175
internet of things	264	sensor technology	172
smart home	241	cloud computing	171
cost efficiency	238	policy	170
security	236	energy	160
platform	233	risk assessment	158
emergency	228	functionality	150
remote network monitoring	204	heart rate	144
human computer interaction(HCI)	202	privacy	136

4) 빈도분석 결과 (2016-2018.04)

최근 3년간(2016년에서 2018년 4월까지)의 연구에서는 순위권에서 밀려났었던 wearable technology가 다시 등장했으며 2011년-2015년에 이어 smart city, smart device, big data와 같은 연구주제가 새롭게 등장한 것을 통해 최첨단 기술과 융·복합된 연구들이 중점적으로 이루어지고 있음을 파악할 수 있다.

표 6. 빈도 분석 결과(2016-2018.04)

Table 6. Result of Frequency Analysis(2016-2018.04)

Word	Freq	Word	Freq
internet of things	929	energy	205
cryptography	445	real time data	200
cloud computing	376	wearable technology	186
medical doctor	354	algorithm	182
security	334	heart rate	182
smart phone	320	remote network monitoring	180
medical prescription	316	platform	179
machine learning	310	policy	177
context awareness	258	personal health system(PHS)	169
smart home	250	privacy	169
feature extraction	246	m-health	167
smart city	244	chronic condition	158
wireless network	231	smart device	157
emergency	218	human computer interaction(HCI)	148
cost efficiency	206	big data	148

4-2.토픽모델링(Topic Modeling)

비정형 데이터로 구성된 논문 초록에서에 잠재되어 있는 주요 토픽(연구주제)을 찾기 위해 토픽모델링을 수행하였다. LDA 기반 토픽모델링을 수행하기 위해서 α 값을 2.0, β 값을 0.001 그리고 샘플링 반복횟수를 1,000번 반복하도록 설정했다. 토픽모델링을 통해 유의미한 결과를 도출하기 위해서는 토픽 수 결정이 매우 중요하다. 그러나 적절한 토픽 수에 대한 통계적 해법은 없다. 토픽 수 결정은 도출된 토픽들의 해석가능성과 타당도 및 연구질문과 관련된 유용성에 따라 좌우되기 때문이다[20].

본 연구에서는 2001년부터 2018년 4월까지 수집된 전체 논문 수 2690편을 대상으로 토픽 수를 4개 부터 15개까지 설정하여 비교 분석한 결과 토픽 수 8개가 해석 상 가장 유의미하고 적절한 것으로 판단되었다. <표 7>은 토픽모델링 분석한 결과 추출된 각 토픽(연구주제)별로 할당확률이 높게 나타난 단어 5개를 순서대로 제시하고, 토픽이 가장 대표적으로 나타난 문서의 비중을 제시하였다. 토픽명은 토픽을 이루고 있는 단어들의 연계성을 고려하여 선정하였다. 논문 2690편의 연구주제 점유율은 AI in healthcare(16%)가 가장 높은 것으로 나타났으며, 그 다음으로 Smart hospital(15%), Healthcare platform(14%), blockchain in healthcare(14%), Smart health data(12%), Mobile healthcare(11%), Wellness care(9%), Cognitive healthcare(9%) 순으로 나타났다.

표 7. 토픽 모델링 결과(2001-2018.04)

Table 7. Result of Topic Modeling(2001-2018.04)

Topic	Word	Doc. (%)
1 AI in healthcare	context awareness, machine learning, feature extraction, heart rate, wearable technology	426 (16%)
2 Smart hospital	wireless network, cloud computing, energy, sensor technology, remote network monitoring	408 (15%)
3 Healthcare platform	internet of things(IoT), smart home, platform, smart city, information and communication technology(ICT)	385 (14%)
4 blockchain in healthcare	cryptography, security, health data, telemedicine, smart card	374 (14%)
5 Smart health data	medical doctor, medical prescription, risk assessment, pharmacy, life style disease	324 (12%)
6 Mobile healthcare	smart phone, personal health system, m-health, nurse, privacy	289 (11%)
7 Wellness care	policy, cost efficiency, chronic condition, quality of life(QoL), functionality	253 (9%)
8 Cognitive healthcare	human computer interaction(HCI), detection, machine learning, assistance, fall	231 (9%)

4-3. 에고 네트워크 분석(Ego Network Analysis)

본 연구에서는 토픽모델링을 통해 추출된 8개의 주요 연구 주제(토픽)를 미시적 관점에서 살펴보기 위해 연구주제별 할당

된 단어들을 대상으로 워드 클라우드와 에고 네트워크 분석)을 시행하였다 에고 네트워크에서 컴포넌트(Component)는 한 그래프 내에서 연결된 그래프를 의미하며, 그래프 전체가 하나인 경우 컴포넌트는 1개 이며, 그래프가 세 개인 경우 컴포넌트는 3개이다. 밀도(Density)는 에고 노드와 연결된 이웃 노드가 얼마나 잘 연결되어 있는지 나타내는 지표[18]로, 모든 노드가 연결되어 있으면 '1'이고 아무런 연결이 없으면 '0'이다.

1) Topic 1. AI in healthcare

연구주제 점유율이 가장 높게 나타난 AI in healthcare(이하 '연구주제 1')는 논문은 전체 논문 2690편 중 426편(16%)에서 나타났다. 2001년부터 2017년 12월까지) 나타난 논문 수는 <그림 2>와 같이 2009년부터 2013년까지 완만한 상승 이후 급격한 증가 추세를 보이고 있다. 본 주제와 관련된 주요 단어를 직관적으로 파악하기 위해 주어진 단어 92개의 할당확률 기준으로 워드 클라우드(Word Cloud)를 <그림 3>와 같이 나타낸 결과 context awareness, machine learning, feature extraction, hart rate, wearable technology의 주요도가 높게 나타났다. 연구주제 1에서 주요도가 높게 나타난 context awareness(상황인지)는 사용자의 행동이나 음성 혹은 사용자를 둘러싼 환경의 변화를 인식하고, 내재된 의도나 요구에 적합하게 환경을 제어하고 서비스를 제공하는 것을 말한다[21]. machine learning(기계학습)은 인공지능의 한 분야로 특정 상황에 대한 꾸준한 학습을 통해 성능을 높이는 것을 말한다. 기계학습에는 지도학습(Supervised Learning)과 비지도 학습(Unsupervised Learning)이 있는데, 에고 네트워크에서 machine learning은 분류(classifier), 예측(prediction)과 연결되어 있기 때문에 주제 1에서는 지도학습을 의미하는 것으로 볼 수 있다. wearable device는 사용자와 24시간 밀착되어 사용자의 상황을 인식하고 다양한 편의를 제공할 수 있기 때문에 IoT 혹은 IoE(Internet of Everything) 시대를 여는데 앞장설 수 있는 기기라고 할 수 있다[22].

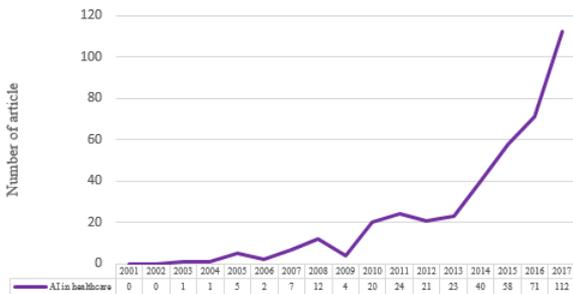


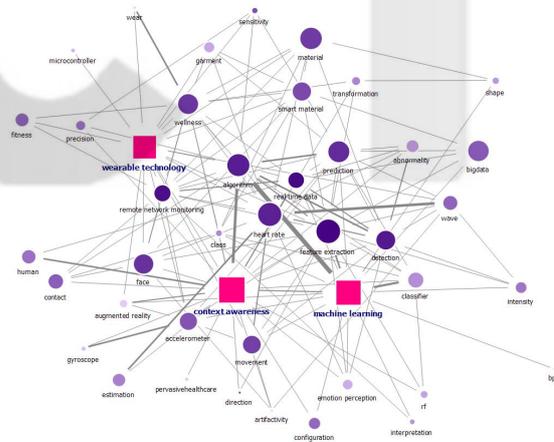
그림 2. 논문 수의 변화(AI in healthcare)
Fig. 2. Changes of the number of article(AI in healthcare)

- 3) 네트워크 중 사각형 노드는 “에고 노드”이며, 토픽에서 할당확률이 높을수록 노드의 크기가 크고, 네트워크 연결중심성(degree centrality)이 높게 나올수록 색이 진해짐. 링크의 굵기는 링크의 가중치에 비례함.
- 4) 2018년도의 경우 4월까지의 논문 발표 수는 다른 연도 1년간의 논문발표 수와 같이 비교하는 것은 무리가 있어 논문 수의 변화는 2017년12월까지 게재된 논문에 한함.



그림 3. 워드 클라우드 결과(AI in healthcare)
Fig. 3. Result of Word Cloud(AI in healthcare)

연구주제 1의 텍스트 네트워크의 구조를 효과적으로 분석하기 위해 주요도가 높게 나타난 context awareness, machine learning, wearable technology 중심으로 에고 네트워크 분석을 <그림 4>과 같이 했다. context awareness는 wearable technology와 machine learning과 연결되어 있으며, human, pervasive healthcare, precision, abnormality 등의 단어와 연결되어 있는 것으로 나타났다. machine learning은 wearable technology와 연결되어 있지 않으며, big data, movement, intensity와 연결되어 있고, real time data, classifier, prediction, detection, emotion perception등의 단어가 context awareness와 같이 연결되어 있는 것으로 나타났다. wearable technology는 remote network monitoring, heart rate, sensitivity, smart material, fitness 등의 단어와 연결되어 있으며, algorithm, feature extraction, augment reality(AR), accelerometer, wellness 등의 단어는 주제 1에서 주요도가 높게 나타난 세 단어에 모두 등장한 것으로 나타났다.



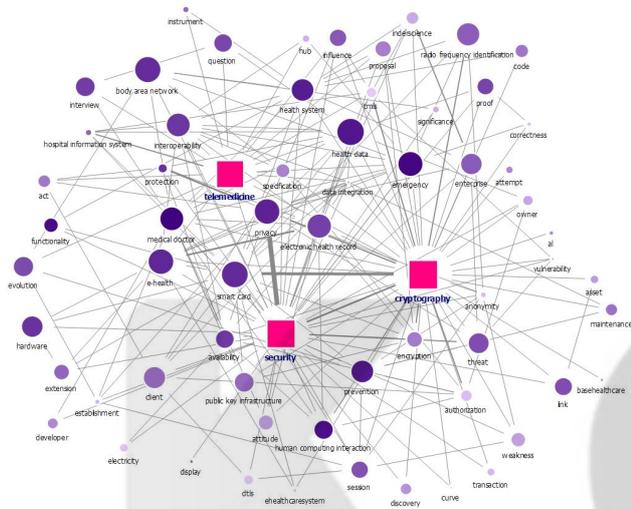
component: 1 node: 43 link: 157 density: 0.174

그림 4. 에고 네트워크 분석(AI in healthcare)
Fig. 4. Ego Network Analysis(AI in healthcare)

2) Topic 2. Smart hospital

Smart hospital(이하 '연구주제 2')의 연구주제 점유율은 전체의 15%(408편)으로 나타났으며, 논문 수는 <그림 5>와 같은 증가 추세를 보이고 있다. 연구주제 2의 주요한 단어를 파악하기 위해 <그림 6>과 같이 단어 91개의 할당확률 기준으로 워드 클라우드로 나타낸 결과 연구주제 2에서는 wireless network, cloud

연구주제 4의 텍스트 네트워크 구조를 효과적으로 분석하기 위해 주요도가 높게 나타난 cryptography, security, telemedicine 중심으로 <그림 13>과 같이 에고 네트워크 분석을 했다. cryptography는 security, telemedicine과 연결되어 있으며, public key infrastructure(PKI), radio frequency identification(RFID), electronic health record(EHR)와 연결되어 있는 것으로 나타났다. security는 body area network(BAN) 등과 연결되어 있으며, cryptography와 같이 연결된 단어는 smart card, human computer interaction(HCI), health system, privacy, prevention, encryption 등이 있는 것으로 나타났다. 연구주제 4의 주요도가 높게 나타난 세 단어와 모두 연결된 단어는 e-health, health data, emergency, specification, medical doctor, availability 등으로 나타났다.



component: 1 node: 65 link: 235 density: 0.113
 그림 13. 에고 네트워크 분석(Blockchain in healthcare)
 Fig. 13. Ego Network Analysis(Blockchain in healthcare)

5) Topic 5. Smart health data

Smart health data(스마트 헬스데이터)는 빠른 속도로 생성되는 대량의 보건의료 데이터(Health data)를 수집하고 분석해서 가치 있는 정보로 추출해 내는 기술로 의료분야의 빅데이터라고 할 수 있다[9]. 스마트 헬스데이터는 가속화되고 있는 인구 고령화와 만성질환의 증가 등으로 인한 의료비 급증으로 치료 중심에서 관리와 예방 중심으로 의료 패러다임이 변화되며, ICT와 헬스케어의 융·복합 기술도입이 증가되면서 급부상했다.

Smart health data(이하 ‘연구주제 5’)의 연구주제 점유율은 전체의 12%(324편)로 나타났으며, 논문 수는 <그림 14>에서 보는 바와 같이 상승과 하락을 반복하고 있는 것으로 나타났다. 연구주제 5의 주요한 단어를 파악하기 위해 <그림 15>과 같이 주어진 단어 87개의 할당확률 기준으로 워드 클라우드를 나타낸 결과 medical doctor, medical prescription, risk assessment, pharmacy, lifestyle disease가 주요도가 높게 나타났다. 연구주제 5

에서 주요도가 높게 나타난 risk assessment(위험 평가)는 위험의 잠재적 손실의 크기와 확률을 측정하는 것으로 위험 관리 처리의 가장 중요한 단계이며, 위험이 확인되어 평가되면 적절한 위험을 다루는 단계들은 계획적이 될 수 있다[25].

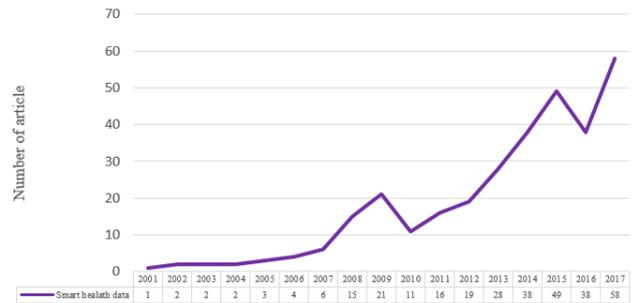
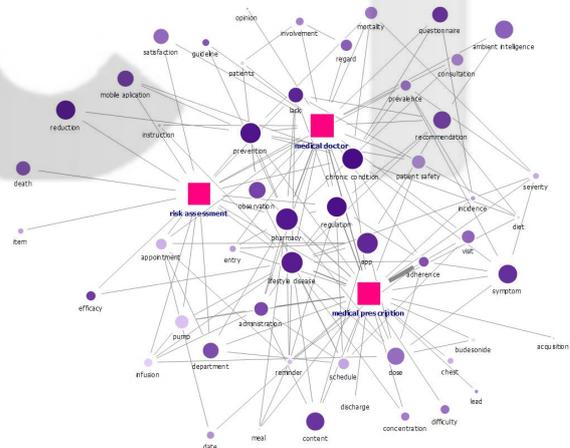


그림 14. 논문 수의 변화(Smart health data)
 Fig. 14. Changes of the number of article(Smart health data)



그림 15. 워드 클라우드 분석(Smart health data)
 Fig. 15. Result of Word Cloud(Smart health data)



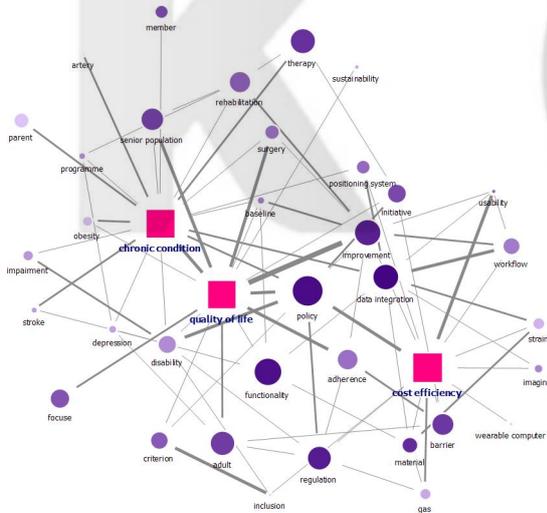
component: 1 node: 55 link: 158 density: 0.106
 그림 16. 에고 네트워크 분석(Smart health data)
 Fig. 16. Ego Network Analysis(Smart health data)

연구주제 5의 텍스트 네트워크 구조를 효과적으로 분석하기 위해 주요도가 높게 나타난 medical doctor, medical prescription, risk assessment 중심으로 <그림 16>과 같이 에고 네트워크를 분석했다. medical doctor는 medical prescription, risk assessment와 연결되어 있으며, ambient intelligence(AmI), app, recommendation 등의 단어와 연결되어 있는 것으로 나타났다.



그림 21. 워드 클라우드 결과(Wellness care)
Fig. 21. Result of Word cloud(Wellness care)

연구주제 7의 텍스트 네트워크 구조를 효과적으로 분석하기 위해 주요도가 높게 나타난 cost efficiency, chronic condition, quality of life(QoL) 중심으로 <그림 22>과 같이 에고 네트워크 분석을 했다. cost efficiency는 chronic condition, quality of life와 연결되어 있지 않으며, regulation, barrier, improvement 등의 단어와 연결되어 있는 것으로 나타났다. chronic condition은 therapy, rehabilitation, data integration 등의 단어와 연결되어 있으며, quality of life와 연결되어 있고 같이 연결된 단어는 senior population, policy, disability 등이 있는 것으로 나타났다. quality of life는 functionality, improvement, adherence 등의 단어와 연결되어 있으며, 연구주제 7의 주요도가 높은 세 단어와 policy가 연결되어 있는 것으로 나타났다.



component: 1 node: 38 link: 86 density: 0.122
그림 22. 에고 네트워크 분석(Wellness care)
Fig. 22. Ego Network Analysis(Wellness care)

8) Topic 8. Cognitive healthcare

Cognitive healthcare(이하 ‘연구주제 8’)의 연구주제 점유율은 9%(231편)으로 나타났으며, 논문 수는<그림 23>에서 보는 바와 같이 연도별로 점진적인 상승추세를 보였고, 2016년을 기점으로 급증한 것으로 나타났다. 연구주제 8의 주요한 단어를 파악하기 위해 <그림 24>과 같이 주어진 단어 70개의 할당확

률을 기준으로 워드 클라우드를 분석한 결과 human computer interaction(HCI), detection, machine learning, assistant가 주요도가 높게 나타났다. 연구주제 8에서 주요도가 높게 나타난 human computer interaction(인간-컴퓨터 상호작용)은 컴퓨터 시스템과 컴퓨터 사용자인 인간 사이의 상호작용을 향상시키기 위한 효과적인 방법을 중점적으로 연구하는 학문 분야로 인간요소(human factor), 인간공학, 산업공학, 인지심리학, 운영체제, 컴퓨터 과학 등 학제 간 융·복합연구가 활발히 진행되고 있다[26].

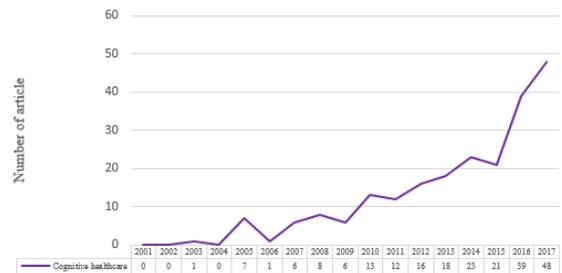
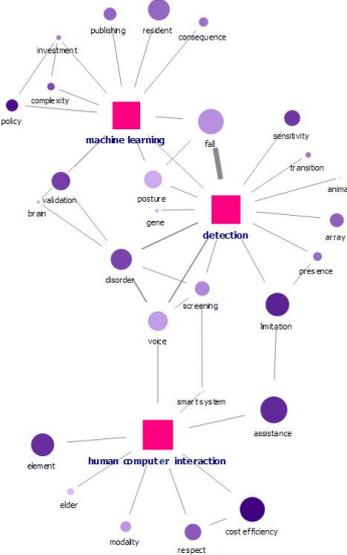


그림 23. 논문 수의 변화(Cognitive healthcare)
Fig. 23. Changes of the number of article(Cognitive healthcare)



그림 24. 워드 클라우드 결과(Cognitive healthcare)
Fig. 24. Result of Word cloud(Cognitive healthcare)

연구주제 8의 텍스트 네트워크 구조를 효과적으로 분석하기 위해 주요도가 높게 나타난 human computer interaction, detection, machine learning 중심으로 <그림 25>과 같이 에고 네트워크 분석을 했다. human computer interaction은 assistance, element, cost efficiency 등의 단어와 연결된 것으로 나타나며, detection과 voice가 연결된 것으로 나타났다. detection은 limitation, sensitivity, disorder, screening 등의 단어와 연결되어 있으며, machine learning과 연결되어 있고 fall, posture가 같이 연결되어 있는 것으로 나타났다. machine learning은 resident, validation, complexity, brain 등의 단어와 연결되어 있는 것으로 나타났다.



component: 1 node: 30 link: 43 density: 0.099

그림 25. 에고 네트워크 분석(Cognitive healthcare)
 Fig. 25. Ego Network Analysis(Cognitive healthcare)

V. 결 론

본 연구의 목적은 스마트 헬스케어산업에서 최첨단 융·복합 기술 발전의 메가트렌드를 형성하고 있는 ‘스마트 헬스케어’ 분야의 최근 주요 연구동향을 파악하고 향후 학제 간 융·복합연구의 전개 방향을 조망해 보기 위한 것이다. 이를 위해 scopus에서 2001년부터 2018년 4월까지 게재된 영문 논문 초록을 추출해서 텍스트 분석, 구간별 빈도분석, 토픽모델링을 통해 8개의 중점 연구주제(토픽)를 도출해 낼 수 있었다. 8개의 주요 연구주제 중 ‘AI in healthcare(16%)가 연구점유율이 가장 높은 것으로 나타났고, 그 다음으로 ‘Smart hospital(15%)’, ‘Healthcare platform(14%)’, ‘blockchain in healthcare(14%)’, ‘Smart health data(12%)’, ‘Mobile healthcare(11%)’, ‘Wellness care(8%)’, ‘Cognitive healthcare(9%)’ 순으로 나타났다. 빈도분석에서는 기간별 중요성을 가지는 연구주제들이 무엇인지 살펴보았다. 그 결과 기간에 상관없이 중요하게 나타난 주제(고빈도 top 10)는 cryptography, medical doctor, medical prescription, smart home으로 나타났다. 이를 통해 암호화, 전문 의료진, 진료행위, 스마트 홈 관련 연구가 지속적으로 중요하게 다루어져 왔음을 알 수 있다. 2001년-2010년간 5년 단위의 기간별 분석에서는 2006년 이후 새롭게 등장한 중점 연구주제들(smart phone, wireless network, platform 등)과 순위 밖으로 밀려난 연구주제들이 있었고 이를 통해 기간별 중점 연구주제의 변화가 있었음을 알 수 있다. 2011년에서 2015년에는 IoT, real time data, feature extraction, m-health 등의 단어가 급상승 하였고, 2016년 1월 WEF에서 ‘4차 산업혁명’ 언급 이후부터 2018년 4월까지의 기간분석에서는 순위권 밖으로 밀려났던 wearable technology가

다시 순위권 안으로 들어왔을 뿐만 아니라 smart device, big data, smart city가 새롭게 중요한 연구주제로 등장하면서 헬스케어산업과 연관산업간의 최첨단 기술들이 융·복합되고 있는 분야가 중점적으로 연구되고 있음을 알 수 있었다.

8개의 중점 연구주제별 에고네트워크 분석을 통해 주제별 연구동향을 파악할 수 있었다. 토픽별 에고 네트워크 분석결과는 ① ‘AI in healthcare’에서는 context awareness, machine learning, wearable technology ② ‘Smart hospital’에서는 wireless network, cloud computing, sensor technology ③ ‘Healthcare platform’에서는 internet of technology, platform, smart home, smart city ④ ‘blockchain in healthcare’에서는 cryptography, security, telemedicine ⑤ ‘Smart health data’에서는 medical doctor, medical prescription, risk assessment ⑥ ‘Mobile healthcare’에서는 smart phone, personal health system, m-health ⑦ ‘Wellness’에서는 cost efficiency, chronic condition, quality of life ⑧ ‘Cognitive healthcare’에서는 human computer interaction, detection, machine learning이 중요한 핵심노드임이 드러났고, 이와 관련된 연구가 활발히 진행되고 있음을 알 수 있었다.

본 연구에서 도출한 ‘스마트 헬스케어’ 관련 8개 중점 연구주제와 연구주제별 에고 네트워크 분석을 통해 파악된 주제별 중점 연구동향들은 향후 이 분야 연구자들이 연구방향을 수립하거나 정부의 기술지원 정책 방향과 전략수립에 중요한 기초자료로 활용할 수 있을 것이다. 본 연구는 논문 초록만을 대상으로 분석하였지만, 국내·외 연구동향, 특허기술 등 다양한 데이터를 기반으로 분석을 한다면 보다 유의미한 시사점을 도출할 수 있을 것으로 기대된다.

참고문헌

[1] K. J. Song, The fourth industrial revolution of Klaus Schwab, *new present, korea*, pp. 24-261, Apr 2016
 [2] J. K. Kim, S. J. Lee, ICT convergence medical industry strategy and implications of major countries, *KIEP World Economy Today*, pp. 1-15, Vol.16, No.23, Jun 2016
 [3] M. J. Kim, Y. Lee, The fourth industrial revolution of Klaus Schwab, *new present*, pp. 24-307, korea, Apr 2018
 [4] M. K. Lee, Trend and implications of health care in the 4th industrial revolution, *weekly KDB Report*, pp.1-13, Jul 2017
 [5] Y. J. Jo, Opportunities to find in the rise of smart healthcare, *KDB Industrial Banking Industry Issue*, pp.43-64, Mar 2016
 [6] Korea Technology and Information Promotion Agency for SMEs, Technology roadmap for SME 2018-2020: *Medical service devices*, korea, 2017
 [7] A. Solanas, C. Patsakis, M. Conti, I. S. Vlachos, V. Ramos, F. Falcone, O. Postolache, P. A. Perez-Martinez, R. D. Pietro, D. N. Perrea, A. Martinez-Balleste, “Smart health:

- A context-aware health paradigm within smart cities”, *IEEE Communications Magazine*, Vol.52, No.8, pp. 74-81, Aug 2014
- [8] M. Y. Kang, D. H. Park, G. S. Kim, Present and Future of Smart healthcare, *Samsung KPMG Economics Issue Monitor*, Vol.79, pp.1-19, Jun 2018
- [9] S. J. Ahn, 2016 Standard-based R&D Roadmap: *Smart health, korea*, pp. 4-339, Apr 2016
- [10] S. H. Kim, “Trends in Medical IT Convergence Technology Research,” *Magazine of the IEEK*, Vol. 43, No. 2, pp. 18-24, Feb 2016
- [11] T. G. Lee, “Smart healthcare and Health-Medical Information System Enforcement Strategies”, *Korean Institute of Information Technology*, Vol. 11, No.1, pp. 41-48, Jun 2013
- [12] I. J. Hwang, M. S. Kang, “Trend and Analysis of electrocardiographic medical device technology for smart healthcare,” *The Institute of Electronics Engineers of Korea*, Vol. 44, No.2, pp.29-35, Feb 2017
- [13] D. W. Kim, G. H. Han, “Recent research trends for responding to security threats in smart medical environment,” *Journal of The Korean Institute of Communication Sciences*, Vol.35, No. 2, pp. 95-99, Feb 2018
- [14] D. M. Blei, “ Probabilistic topic models”, *Communication of the ACM*, Vol. 55, No. 4, pp. 77-84, Apr 2012
- [15] D. M. Blei, A. Y. Ng, M. I. Jordan, “Latent Dirichlet Allocation”, *Journal of Machine Learning Research* Vol. 3, No. 4-5, pp. 993-1022, may 2003
- [16] C. S. KIM, S. J. Choi, K. Y. Kwahk, “Investigation of Research Trends in Information Systems Domain Using Topic Modeling and Time Series Regression Analysis”, *Journal of Digital Contents Society*, Vol. 18, No. 6, pp.1143-1150, Oct 2017
- [17] C. S. Park, J. W. Jeong, “Text Network Analysis: socio-cognitive network analysis to understand shared meaning among policy stakeholders,” *Journal of Governmental Studies*, Vol, 19, No. 2, pp. 73-108, Aug 2013
- [18] S. S. Lee, *Network Analysis Methodology*, 4th ed, nonhyeong, pp.5-370, 2012
- [19] G. H. Jeong, A study of foresight method based on textmining and complexity network analysis, *Korea Institute of Science & Technology Evaluation and Planning*, pp. 1-138, Dec 2010
- [20] C. H. Nahm, “An Illustrative Application of Topic Modeling Method to a Farmer’s Diary,” *Institute of Cross Cultural Studies*, Vol. 22, No. 1, pp. 89-135, 2016
- [21] telecommunications Technology Association[internet]. Available: <http://terms.tta.or.kr/main.do>
- [22] I. H. Hwang, S. G. Oh, “Trends of Situation Awareness Using Wearable Devices,” *Telecommunication Review*, Vol. 24, No. 4, pp. 492-499, Jul 2014
- [23] K. M, Park, W. H. Seok, K. H. Lee, Sensor industry and major prospective sensor market and technology trend: Focused on Bio(Medical), *Mobile and Automotive Fields, ETRI Creative Opinion*, Issue Report 2015-02, May 2015
- [24] H. J. Choi, Application of blockchain in the healthcare industry, *KHIDI Brief*, Vol.236, may 2017
- [25] Korean dictionary[internet]. Available:<https://opendict.korean.go.kr/main>
- [26] HRD term dictionary[internet]. Available: <https://terms.naver.com/entry.nhn?docId=2178563&cid=51072&categoryId=51072>



윤지은 (Jee-Eun Yoon)

2012년 ~ 2014년
: 한양대학교 경영전문대학원
글로벌 의료경영MBA(경영학석사)

2016년 ~ 현재 : 한양대학교 일반대학원
경영학과 박사과정 수료

※관심분야 : 스마트 헬스케어(Smart Healthcare), 텍스트 마이닝(Text Mining), 토픽모델링(Topic Modeling), 네트워크 분석(Network Analysis), 빅데이터, 디지털 비즈니스(Digital Business) 등



서창진 (Chang-Jin Suh)

1994년 : Vanderbilt Univ. 경제학 박사

1994년~1999년: 한국의료관리연구원 연구위원
1999년~2008년: 한국보건산업진흥원 전문위원
2008년~현재 : 한양대학교 경영학부 교수
※관심분야 : 보건경제, 의료경영, 의료정책, 보건의료산업정책

КСІ