

Exploring the Precision Medicine Research Trend using Topic Modeling

토픽 모델링을 활용한 정밀의료 연구동향 탐색

Su Hyun Ahn¹, Jun Hyeok Ko², Sang Jun Lee³

안수현¹, 고준혁², 이상준³

¹ Professor, College of General Education, Semyung University, South Korea,
tanny10@semyung.ac.kr

² Student, Department of Business Administration, Semyung University, South Korea,
kojunhuk@semyung.ac.kr

³ Professor, College of General Education, Semyung University, South Korea,
leesangjun@semyung.ac.kr

Corresponding author: Sang Jun Lee

Abstract: Precision medicine is a technology that provides optimal and customized healthcare services by analyzing an individual's genetic information and life information. This study aims to investigate the current research trends in precision medicine and provide recommendations for the field's future direction. Therefore, after data collection and preprocessing, major keyword analysis and topic modeling were conducted on domestic media reports. As a result of the analysis, Topic 1 is "Industrial innovation and technology-driven job creation", Topic 2 is "Innovation of precision medical services and utilization of big data", Topic 3 is "Advancement of precision medicine through research and innovation", Topic 4 is "Precision medical data management using big data and cloud", and Topic 5 is "Advancement of genomic and biological research", Topic 6 is "Business and innovation through biohealth and convergence technologies," Topic 7 is "Competition and collaboration in the global marketplace," Topic 8 is "Collaboration between hospitals centered on precision medicine and data technologies," and Topic 9 is "Research and development for drug discovery and personalized care." The findings are expected to provide new insights into various fields by identifying key topics and trends in precision medicine.

Keywords: Precision Medicine, Text Mining, Topic Modeling, Research Trend

요약: 정밀의료는 개인의 유전정보, 생활정보 등을 분석하여 최적의 맞춤형 헬스케어 서비스를 제공하는 기술이다. 본 연구의 목적은 정밀의료 분야의 연구동향을 탐색하고, 이를 통해 정밀의료 분야의 발전방향을 제시하는 것이다. 이에 국내 언론보도 기사자료를 대상으로 데이터 수집 및 전처리 이후 주요 키워드 분석 및 토픽 모델링을 수행하였다. 분석 결과 토픽 1은 “산업 혁신과 기술 중심의 일자리 창출”, 토픽 2는 “정밀의료 서비스 혁신과 빅데이터 활용”, 토픽 3은 “연구와 혁신을 통한 정밀의료 발전”, 토픽 4는 “빅데이터와 클라우드를 활용한 정밀의료 데이터 관리”, 토픽 5는 “유전체와 생물학적 연구의 발전”, 토픽 6은 “바이오헬스와 융복합 기술을 통한 사업과 혁신”, 토픽 7은 “글로벌 시장에서의 경쟁과 협력 관계”, 토픽 8은 “정밀의료와 데이터 기술을 중심으로 하는 병원 간 협력”, 토픽 9는 “신약

Received: September 13, 2023; 1st Review Result: October 14, 2023; 2nd Review Result: November 18, 2023
Accepted: December 26, 2023

개발과 맞춤형 치료를 위한 연구와 개발”로 나타났다. 연구결과를 통해 정밀의료의 핵심 주제 및 동향을 파악함으로써 다양한 분야에 새로운 통찰력을 제공할 것으로 기대한다.

핵심어: 정밀의료, 텍스트마이닝, 토픽 모델링, 연구동향

1. 서론

2015년 미국 오바마 대통령이 기존의 의료체계를 개선하고 혁신적인 의료방식을 도입하기 위한 계획으로 “정밀의료 발전계획(Precision Medicine Initiative, PMI)”을 발표하면서 정밀의료는 차세대 패러다임으로 급부상하였다. 이후 국제적으로 널리 확산되면서 한국 뿐만 아니라 유럽과 중국, 일본 등 다양한 국가에서도 정밀의료 프로젝트 및 발전계획이 등장했다[1]. 이러한 발전과정은 정밀의료의 단순한 의료기술의 일부가 아니라 국가적 차원에서의 중요한 전략적 분야로 부상함을 보여주며, 환자 중심의 의료 제공과 새로운 기술의 도입을 통해 의료혁신을 이루어 나가는 전환점이라 볼 수 있다. 정밀의료는 대규모 유전체 정보분석을 통해 미리 예방하고 선제적으로 개별화된 헬스케어 서비스를 제공하는 방향으로 발전하고 있다. 최근에는 각종 질환의 진단과 치료에 있어 환자들마다 서로 다른 접근과 치료가 필요하다는 점이 밝혀짐에 따라 정밀의료의 중요성이 강조되고 있다[2].

정밀의료(precision medicine)는 환자마다 다른 유전적, 환경적 요인, 질병 경력, 생활습관 등을 미리 파악하여 환자에게 맞는 약과 적절한 용량을 사용하여 최적의 치료를 제공하는 것을 의미한다[3]. 정밀의료 서비스는 유전체 정보, 진료 및 임상정보, 생활습관 정보 등을 통합하여 환자 개인의 특성에 맞는 맞춤형 의료서비스를 제공함으로써 진료의 정확도를 높이고 치료 효과를 최적화할 수 있다. 정밀의료 서비스는 맞춤형료(personalized medicine)와 혼용해서 사용하기도 하는데[4], 맞춤형료의 경우 특정한 개인을 위해 예방법 및 치료법이 개발되고 있는 것처럼 오해소지가 있어서 최근 미국국립연구회의(National Research Council)에서는 정밀의료 서비스로 사용하는 것을 권장하고 있다[5]. 정밀의료는 개인의 유전정보, 질병정보, 생활정보 등을 종합적으로 분석하여 보다 정교하게 환자를 분류하고 이를 기반으로 특화된 치료법을 선택한다는 측면에서 맞춤형료의 개념을 구현하고 있다[6].

현대 의료 분야에서 과학과 기술의 급속한 발전으로 인해 정밀의료는 혁신적인 기술과 방법을 도입하면서 급격한 변화를 겪고 있다. 정밀의료 분야는 인공지능, 빅데이터, 바이오기술 등 다양한 학문 분야의 융합이 필요한 분야이고 다양한 주제와 이슈들이 존재하기 때문에 최신 동향을 파악하여 앞으로의 방향성을 살펴볼 필요가 있다. 연구 동향은 다양한 측면에서 발전하고 있기 때문에 이를 종합적으로 이해하고 핵심 내용을 파악하여 주요 이슈와 트렌드에 민감하게 대응하는 것이 중요하다. 그러나 정밀의료는 다양한 분야의 연구가 교차되는 속성이 있기 때문에 주제 간의 연결고리를 찾아내어 각 분야 간의 상호작용을 이해하는데 어려움이 있다.

최근 빅데이터 분석을 활용하여 연구동향을 파악하는 연구가 많아지고 있는데, 토픽 모델링을 활용한 연구는 이러한 문제에 대응하기 적합하고 복잡한 연구 분야에서 동시에 다양한 주제를 이해하는데 유용하다. 토픽 모델링은 주제를 뜻하는 토픽(topic)과 모형화를 뜻하는 모델링(modeling)이 결합된 개념으로 구조화되지 않은 대량의 텍스트에서 숨겨진 주제 구조를 발견하기 위한 통계적 추론 알고리즘이다[7]. 또한

비구조적인 텍스트 데이터는 다양한 주제를 다루며 이를 구조화하여 파악하기 어려운데, 토픽 모델링은 텍스트를 특정 주제와 연관된 그룹으로 구조화하여 정보를 효율적으로 관리하고 해석할 수 있게 해줌으로써 연구동향의 주요 주제들을 명확하게 추출할 수 있다. 실제 선행연구를 살펴보면 과학기술 연구동향[8], 교통수단 연구동향[9], 생물정보학 연구동향[10], 컴퓨터 언어 연구동향[11] 등 다양한 분야의 연구동향을 탐색하기 위해 토픽 모델링이 활용되는 것을 볼 수 있다.

정밀의료 영역은 앞으로 다가올 미래 시대에 가장 큰 변화가 예상되는 분야가 될 것이라는 전망과 함께 급변하는 정보화 사회 적응과 지속가능한 의료 서비스를 위하여 새로운 논의의 필요성이 제기되고 있다. 이러한 시대적 흐름 속에서 본 연구는 토픽 모델링을 활용한 정밀의료 연구동향 탐색을 수행하고자 한다. 보다 구체적으로 뉴스 기사를 수집 분석하여 정밀의료 분야에서 어떠한 주요 키워드들이 다루어지고 있는지 살펴보고, 토픽 모델링을 통해 주요 토픽들을 도출하여 각 토픽의 핵심내용과 연결점을 해석함으로써 정밀의료의 다양한 측면을 종합적으로 탐색한다. 마지막으로 분석결과를 바탕으로 정밀의료 분야의 미래에 대한 지향성과 발전 방향성을 제시하여 다양한 분야의 연구자들에게 유용한 정보를 제공하고자 한다. 본 연구를 통해 정밀의료의 핵심 주제 및 동향을 파악함으로써 현상을 깊이 이해하고 미래발전의 방향성을 제시할 수 있을 것이다.

2. 연구방법

2.1 데이터 수집

정밀의료 분야는 다학제적인 성격을 가지고 있어서 다양한 분야의 전문가와 대중들의 시각을 모두 고려한 보다 포괄적이고 다면적인 연구동향을 파악해야 한다. 이에 본 연구는 정밀의료 연구동향 탐색을 위한 원데이터를 학술지 논문이 아닌 뉴스 기사를 활용하였다. 학술지 논문은 전문가들을 대상으로 하기 때문에 기술적이고 전문적인 언어로 작성되었고, 출판 및 피어 리뷰 과정이 길어 신속한 정보 반영이 어렵다고 판단했다. 반면 뉴스기사는 일반 독자들을 대상으로 하며, 간단하고 이해하기 쉬운 언어로 작성되었고, 실시간으로 새로운 연구 동향과 사건들을 다루기 때문에 기술과 정책, 윤리 등 다양한 측면에서 빠르게 변화하는 정밀의료 분야를 탐색하기에 적합하다. 정밀의료는 다양한 분야의 연구 동향이 상호작용하며 발전하는 복합적인 주제이기 때문에, 뉴스 기사를 활용하여 광범위한 토픽을 파악하고, 이를 종합적으로 분석하여 포괄적인 결과도출을 목표로 하였다.

이에 분석에 사용한 데이터는 한국언론진흥재단이 제공하는 신문기사 검색 서비스인 빅카인즈(Bigkinds)에서 수집하였다[12]. 빅카인즈 서비스는 뉴스 검색 서비스인 KINDS (Korea Integrated News Database System)에서 진화한 새로운 뉴스 분석 서비스로서 1990년부터 시작한 KINDS 서비스를 통하여 주요 일간지, 방송 등 다양한 뉴스를 축적하는 검색 서비스를 제공하고 있다. 또한 종합일간지, 경제지, 지역일간지, 방송사 등과 같은 다양한 언론사로부터 수집한 뉴스로 구성된 통합 데이터베이스에 빅데이터 분석 기술을 접목하여 만든 새로운 뉴스 분석 서비스이다.

데이터 수집 대상 언론사는 1) 전국 일간지 11개(경향신문, 국민일보, 내일신문, 동아일보, 문화일보, 서울신문, 세계일보, 조선일보, 중앙일보, 한겨레, 한국일보), 2) 경제 일간지 8개(매일경제, 머니투데이, 서울경제, 아시아경제, 아주경제, 파이낸셜뉴스,

한국경제, 헤럴드경제), 3) 지역 일간지 28개(강원도민일보, 강원일보, 경기일보, 경남도민일보, 경남신문, 경상일보, 경인일보, 광주매일신문, 광주일보, 국제신문, 대구일보, 대전일보, 매일신문, 무등일보, 부산일보, 영남일보, 울산매일, 전남일보, 전북도민일보, 전북일보, 제민일보, 중도일보, 중부매일, 중부일보, 충북일보, 충청일보, 충청투데이, 한라일보), 4) 방송사 5개(KBS, MBC, OBS, SBS, YTN), 5) 전문지 2개(디지털타임스, 전자신문)의 기사내용을 분석 대상으로 선정하였다.

[그림 1]에서 보는 바와 같이 검색 키워드는 “정밀의료”로 한정하였고, 검색기간은 2013년 1월 1일부터 2022년 12월 31일까지 10년으로 제한하였다. 수집한 기사 자료는 Microsoft Excel 프로그램을 이용하여 하나의 파일로 통합하고, 날짜, 언론사명, 기사제목, 기사분류, 키워드, 본문 내용으로 분류하여 저장하였다. 수집한 자료는 2013년 3건, 2014년 27건, 2015년 34건, 2016년 642건, 2017년 788건, 2018년 989건, 2019년 789건, 2020년 793건, 2021년 952건, 2022년 878건으로 총 5,895건이었다.



[그림 1] 빅카인즈 데이터 수집 과정

[Fig. 1] Bigkinds Data Collection Process

2.2 데이터 분석

본 연구는 정밀의료의 연구동향을 파악하기 위해 빅데이터 분석도구인 R을 활용하여 분석하였다. 빅카인즈에서 수집한 데이터는 비정형 텍스트 데이터이기 때문에 [그림 2]와 같이 전처리를 통해 분석 가능한 형태로 가공하는 절차가 필요하다. 한 토픽에 똑같은 문서가 여러 개 들어 있으면 오류가 생기기 때문에 중복 문서를 제거하고, 3개 키워드 이하로 작성된 짧은 문서는 제거하였다. 이후 문서의 주제는 명사로 결정되기 때문에 중복 사용된 키워드를 제거하면서 명사를 추출하였다. 마지막으로 불용어(들이, 하게,

만큼 등)를 제거하고, 유의어(우리나라=한국, 복지부=보건복지부, 미래부=미래창조과학부 등)를 수정하였다. 또한 형태소 사전에 포함되어 있지 않지만 본 분석에서 의미 있는 것으로 판단되는 키워드(오믹스, 액체생검, 마크로젠 등)는 사용자 사전에 추가하였다[13].

```

1 raw_news_comment <- read_csv("bigkinds.csv") %>% mutate(id = row_number())
2 new_words <- data.frame(term = c("정밀의료", "오믹스", "액체생검"), tag = "ncn")
3 buildDictionary(ext_dic = c("woorimalsam", "insighter"), user_dic = new_words,
4                 replace_usr_dic = T, category_dic_nms = "all")
5 get_dictionary("user_dic")
6 stopword <- read_csv("stopword.csv")
7 news_comment <- raw_news_comment %>%
8   mutate(reply = str_replace_all(reply, "[^가-힣]", " "),
9          reply = str_squish(reply)) %>%
10  distinct(reply, .keep_all = T) %>%
11  filter(str_count(reply, boundary("word")) >= 3)
12 comment <- news_comment %>%
13  unnest_tokens(input = reply, output = word, token = extractNoun, drop = F) %>%
14  filter(str_count(word) > 1) %>% filter(!word %in% stopword$word) %>%
15  mutate(word = recode(word, "우리나라" = "한국")) %>% group_by(id) %>%
16  distinct(word, .keep_all = T) %>% ungroup() %>% select(id, word)

```

[그림 2] 데이터 전처리 과정

[Fig. 2] Data Preprocessing Process

데이터 전처리 이후 본격적으로 토픽 모델링을 실시하였는데, 토픽 모델링은 문서행렬(Document-Term Matrix, DTM)을 기반으로 문서에 잠재되어 있다고 가정된 토픽의 등장확률을 추정하는 통계적 텍스트 처리기법이다. 이러한 방법은 뉴스기사 간의 주제 간 상호 연결성을 고려하여 토픽을 추출하여 연구동향 사이의 상호작용 및 의미있는 패턴을 파악할 수 있다. 연구자들은 가장 널리 알려진 토픽모형인 잠재적 디리클레 할당(Latent Dirichlet Allocation, LDA)모형을 사용하였다[14]. LDA 토픽 모델링은 확률 기반의 모델링 기법을 통해 방대한 양의 문서 데이터를 분석함으로써 문서 내에 어떤 토픽이, 어떤 비율로 구성되어 있는지 분석한다. 또한 토픽별로 어떤 키워드가 구성되었는지 정보를 제공하기 때문에 키워드 조합을 통해 인사이트를 도출하는데 효과적인 장점이 있기 때문이다[15]. 이후 키워드가 토픽에 등장할 확률인 베타(beta)를 산출하고, 문서가 각 토픽에 등장할 확률인 감마(gamma)를 산출하여 주요 키워드와 일문을 함께 살펴보면서 토픽의 특징을 탐색하였다[16].

3. 연구결과

3.1 주요 키워드 분석

2013년부터 2022년까지 10년 동안 게재된 국내 언론보도 기사 5,895건의 내용에서 명사 형태소를 추출하여 데이터 전처리 작업을 거친 후 최종 선정한 주요 키워드는 [표 1]과 같다. 정밀의료(1,441회), 빅데이터(1,387회), 유전체(1,234회)가 가장 많이 나타났으며, 미국(899회), 유전자(891회), 연구(843회), 한국(839회), 인공지능(778회), 차세대(695회), 사업(532회)의 순으로 나타났다.

[표 1] 주요 키워드와 출현횟수

[Table 1] Main Keywords and Frequency

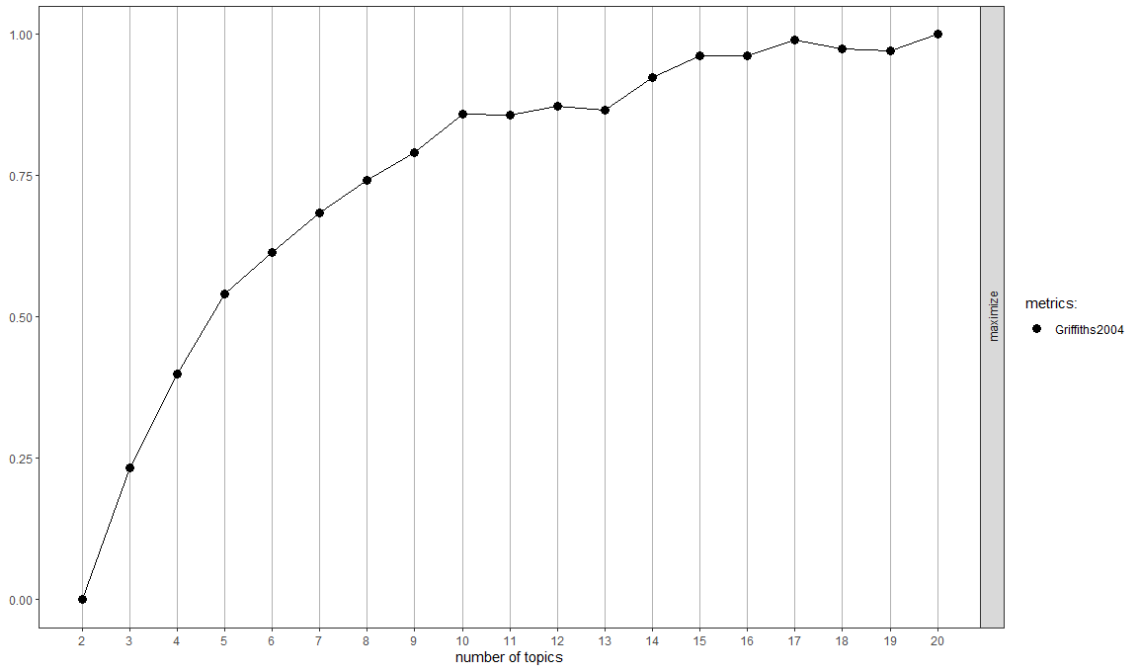
번호	키워드	출현횟수	번호	키워드	출현횟수
1	정밀의료	1,441	26	헬스케어	301
2	빅데이터	1,387	27	보건복지부	285
3	유전체	1,234	28	연구소	273
4	미국	899	29	위원회	268
5	유전자	891	30	항암제	253
6	연구	843	31	의료기기	251
7	한국	839	32	의료진	251
8	인공지능	778	33	관리	247
9	차세대	695	34	서울대병원	246
10	사업	532	35	클러스터	240
11	치료	477	36	마يو헬스	235
12	치료제	439	37	네트워크	234
13	관계	414	38	환자	230
14	전문가	395	39	과학	228
15	서울	389	40	왓슨	220
16	강원	380	41	건강	218
17	일자리	375	42	충북	217
18	경쟁력	366	43	기술	215
19	클라우드	341	44	활성화	215
20	기업	338	45	한국인	213
21	과기정통부	337	46	융복합	212
22	의료	318	47	중국	212
23	의약품	318	48	전문	208
24	개발	305	49	개념	197
25	산업	302	50	업무협약	190

3.2 토픽 모델링

토픽 모델링의 구조는 토픽의 수에 따라 달라진다. 연구자들은 최적의 모형을 결정하기 위해서 토픽의 개수를 2개부터 20개까지 늘려가면서 주요 키워드가 토픽을 잘 대표하는지, 문서가 비슷한 내용끼리 잘 분류되었는지를 판단하였고, [그림 3]과 같이 하이퍼파라미터 튜닝(hyperparameter tuning)을 탐색하면서 토픽을 최종 9개로 설정하였다. 본 연구에 사용한 Griffiths2004는 Griffiths와 Steyvers(2004)가 제안한 복잡도 지표로서 다양한 연구에서 토픽 모델의 성능을 측정하는데 활용되고 있다[17].

연구자들은 토픽의 수와 토픽의 이름을 설정한 후 토픽명이 토픽별 대표 키워드를 잘 포괄하는지에 관해 3차례 회의를 통해 내용 타당도 검증을 실시하였다. 9개 핵심 토픽에 대해 상위 20개 주요 키워드를 가중치 순으로 정리하면 다음 [표 2]와 같다. 분석 결과 토픽 1은 “산업 혁신과 기술 중심의 일자리 창출”, 토픽 2는 “정밀의료 서비스 혁신과 빅데이터 활용”, 토픽 3은 “연구와 혁신을 통한 정밀의료 발전”, 토픽 4는 “빅데이터와 클라우드를 활용한 정밀의료 데이터 관리”, 토픽 5는 “유전체와 생물학적 연구의 발전”, 토픽 6은 “마يو헬스와 융복합 기술을 통한 사업과 혁신”, 토픽 7은 “글로벌 시장에서의 경쟁과 협력 관계”, 토픽 8은 “정밀의료와 데이터 기술을 중심으로 하는 병원 간 협력”,

토픽 9는 “신약 개발과 맞춤형 치료를 위한 연구와 개발”로 나타났다.



[그림 3] 하이퍼파라미터 튜닝

[Fig. 3] Hyperparameter Tuning

[표 2] 토픽 모델링

[Table 2] Topic Modeling

토픽	주요 키워드
토픽 1	일자리, 인공지능, 과학, 기술, 자동차, 정밀의료, 주행차, 경쟁력, 스마트시티, 자율주행차, 반도체, 산업, 중소기업, 뉴딜, 기업, 전남, 광주, 세종, 신산업, 혁신성장
토픽 2	인공지능, 서울대병원, 의료진, 빅데이터, 왓슨, 전문, 대장암, 서울, 길병원, 암센터, 치료, 가천대, 의료서비스, 미국, 서울아산병원, 환자, 생활, 다학제, 건양대병원, 의료비
토픽 3	연구, 정밀의료, 개발, 관리, 과기정통부, 건강, 연구소, 위원회, 헬스케어, 감염병, 보건복지부, 코호트, 국립암센터, 대학원, 경기도, 과학자, 오믹스, 고양, 의료기기, 보험
토픽 4	빅데이터, 강원, 클라우드, 클러스터, 연세, 춘천, 사업단, 의료기관, 원주, 표준화, 고려대의료원, 규제자유특구, 연세의료원, 더존비즈온, 의료데이터, 생태계, 정밀의료, 대학교, 의료원, 간담회
토픽 5	유전체, 유전자, 차세대, 미국, 연구팀, 단백질, 암세포, 생검, 투자, 유방암, 마크로젠, 유럽, 정확, 코스닥, 액체생검, 발병, 한국인, 검사, 바이오마커, 연구진
토픽 6	사업, 바이오헬스, 충북, 융복합, 의료기기, 전문가, 지자체, 후보, 화장품, 스타트업, 교육, 바이오산업, 충주, 일자리, 오송, 대학교, 온라인, 국가산단, 클러스터, 방사선
토픽 7	한국, 미국, 관계, 빅데이터, 서울, 전문가, 중국, 경쟁력, 계몽, 울산, 일본, 헬스케어, 서울대, 영국, 한국인, 산업혁명, 바이오메디컬, 정밀의학, 상용화, 연구소
토픽 8	의료, 사업, 기업, 업무협약, 정밀, 정밀의료, 지원, 네트워크, 병원, 데이터, 공동연구, 산업, 고도화, 의생명, 센터, 환자, 활성화, 평가, 참여, 정보
토픽 9	정밀의료, 치료제, 치료, 의약품, 차세대, 항암제, 유전체, 삼성서울병원, 보건복지부, 신약개발, 제약, 임상시험, 유전자, 부작용, 환자, 맞춤형, 화순전남대병원, 가능성, 식약처, 신테카바이오

3.2.1 토픽 1: 산업 혁신과 기술 중심의 일자리 창출

토픽 1은 정밀의료 분야에서의 산업 혁신과 기술 중심의 일자리 창출에 대한 깊은 이해를 제공한다. 현대 사회에서는 기술 중심의 일자리가 더욱 중요해지고 있으며, 특히 정밀의료 분야에서는 인공지능, 반도체 등과 같은 첨단 기술의 도입이 더욱 두드러지고 있다. 이에 따라 토픽 1은 정밀의료 산업에서의 혁신적 기술 동향을 분석하고, 이로 인한 일자리 창출에 대한 전망과 영향을 조망한다. 정밀의료뿐만 아니라 산업 혁신의 핵심적 측면인 스마트시티, 자율주행차, 그리고 혁신성장을 함께 고려하여 미래 지향적인 연구 방향성을 제시한다.

3.2.2 토픽 2: 정밀의료 서비스 혁신과 빅데이터 활용

토픽 2는 정밀의료 분야의 의료 서비스 혁신과 빅데이터 활용에 대한 체계적인 분석을 제공한다. 정밀의료 서비스 혁신은 환자 중심의 의료 서비스 제공과 연관되어 있으며, 빅데이터는 의료 데이터의 증가로 인한 효과적인 관리와 분석을 지원한다. 토픽 2는 의료 서비스 혁신의 주요 동향을 규명하고, 빅데이터 기술이 어떻게 의료 서비스의 효율성을 향상시키는지에 대한 학술적인 해석을 제공한다. 또한, 빅데이터 활용에 따른 환자 치료 및 대응 등 의료서비스의 변화와 확장 가능성에 중점으로 두고 있다.

3.2.3 토픽 3: 연구와 혁신을 통한 정밀의료 발전

토픽 3은 정밀의료 분야에서의 연구와 혁신의 중요성을 강조한다. 연구와 혁신은 정밀의료의 핵심 가치 중 하나로 새로운 치료법, 의료 기기, 진단 방법 등의 개발을 촉진한다. 토픽 3은 정밀의료 연구의 현재 동향과 미래 전망을 탐구하며, 연구와 혁신이 정밀의료 발전에 어떻게 기여하는지에 대한 학술적 해석을 제공한다. 또한, 정밀의료 연구에서의 다양한 분야의 상호작용과 협력관계를 고찰하여 효과적인 연구 생태계를 조성하는 방안에 대해 논의한다.

3.2.4 토픽 4: 빅데이터와 클라우드를 활용한 정밀의료 데이터 관리

토픽 4는 정밀의료 데이터의 증가로 인한 데이터 관리의 중요성을 강조하며, 빅데이터와 클라우드 기술을 활용한 의료 데이터 관리의 최신 동향을 보여준다. 빅데이터와 클라우드를 의료 분야에 효과적으로 접목시키는 방법과 이로 인한 의료 데이터의 관리 및 활용에 대한 방향성을 제시한다. 의료기관 간의 협업과 표준화된 데이터 활용, 클라우드를 통해 데이터를 안전하게 관리하고 실시간으로 접근 가능하게 하는 정밀한 의료 서비스를 제공하는데 초점을 맞추고 있다.

3.2.5 토픽 5: 유전체와 생물학적 연구의 발전

토픽 5는 유전체 분석과 생물학적 연구의 최신 동향을 탐구하며, 특히 유전체 정보를 기반으로 한 맞춤형 치료 및 예방 전략에 대한 연구에 주목한다. 학술적으로 유전체 분석과 생물학적 연구가 어떻게 정밀 의료의 발전에 기여하는지를 논의하고, 이를 토대로 미래 의료의 방향성을 제시한다. 유전자 및 단백질 연구를 활용하여 암세포의 정확한 진단과 발병 예측을 위한 미래 지향적인 연구 및 투자를 나타내는데, 이는 맞춤형 정밀의료 서비스의 발전 가능성을 제시하고 있다. 유전체와 생물학적 연구는 질병 원인과 예방, 정밀진단과 맞춤형 치료에 대한 연구를 촉진한다.

3.2.6 토픽 6: 바이오헬스와 융복합 기술을 통한 사업과 혁신

토픽 6은 바이오헬스와 융복합 기술이 정밀의료 분야에서 어떻게 사업 모델을 혁신하고 기술 개발에 기여하는지를 학술적으로 분석한다. 바이오헬스와 산업 간의 융복합 기술은 현대 정밀의료 분야에서의 사업모델과 혁신적인 기술개발을 이끌어내고 있다. 바이오헬스와 융복합 기술의 경계를 넘어서는 협력과 혁신이 어떻게 이루어지고 있는지를 심층적으로 조명하며, 이를 통해 의료 분야에서의 미래 지향적 사업 및 기술 발전 방향을 논의한다. 특히 지자체와 협업하여 지역 바이오산업 생태계를 육성하고 지역의 경제적 발전을 이끌어내는데 초점을 맞추고 있다.

3.2.7 토픽 7: 글로벌 시장에서의 경쟁과 협력 관계

토픽 7은 다양한 국가 간의 협력과 경쟁, 특히 글로벌 시장에서의 트렌드와 이슈에 중점을 두고, 국가 간 연구 협력과 산업 협력이 정밀의료 분야에서 어떻게 진행되고 있는지에 대한 관점을 제공한다. 정밀의료는 글로벌 시장에서의 경쟁과 협력이 더욱 중요해지고 국제적인 시각에서도 큰 주목을 받고 있다. 글로벌 시장에서의 경쟁과 협력이 어떻게 정밀의료 기술과 산업의 발전에 영향을 미치는지에 대한 학술적 해석을 통해 전략적인 시사점을 도출한다.

3.2.8 토픽 8: 정밀의료와 데이터 기술을 중심으로 하는 병원 간 협력

토픽 8은 정밀의료 분야에서 병원 간 협력의 중요성과 협업을 위한 데이터 기술의 활용에 대한 내용을 보여준다. 병원 간의 협력과 데이터 기술의 중요성은 계속해서 증가하고 있다. 병원 간의 데이터 공유와 협업이 환자 진료에 어떻게 기여하는지에 대한 깊은 이해를 제시하고, 병원 간 협력과 데이터 기술이 어떻게 환자 중심의 의료 서비스를 높이고 개선할 수 있는지에 대한 전문적인 의견을 제공한다. 정밀의료 분야에서 협업과 데이터 기반의 평가를 통해 의료 시스템을 활성화하고 향상시키는 데 초점을 맞추고 있다.

3.2.9 토픽 9: 신약 개발과 맞춤형 치료를 위한 연구와 개발

토픽9는 신약 개발과 맞춤형 치료가 정밀의료 분야에서 중요하다는 것을 보여준다. 최신 연구 동향을 기반으로 신약 개발과 맞춤형 치료에 대한 내용을 제공하며, 유전자와 환자 데이터를 기반으로 한 맞춤형 치료 전략의 현재와 미래에 대해 다룬다. 또한, 신약 개발과 맞춤형 치료의 도입이 어떻게 진료 효과를 향상시키고 환자의 삶의 질을 향상시키는지에 대한 학술적인 이해를 제시한다. 이는 정밀의료 분야에서의 지속적인 발전을 위해 환자의 유전체 정보를 기반으로 한 맞춤형 치료 및 신약의 개발과정에서의 부작용 최소화에 초점을 맞추고 있다.

4. 결론 및 시사점

본 연구는 정밀의료 연구동향을 탐색하기 위해 국내 언론보도 기사를 대상으로 토픽 모델링을 실시하였다. 9개의 토픽으로 구분된 정밀의료 관련 동향들을 종합적으로 살펴보면 현대 정밀의료 분야에서의 지속적인 혁신과 발전의 모습이 뚜렷하게 드러나고, 첨단 기술의 도입과 데이터 기술의 활용이 중심이 되어 미래 지향적인 발전 방향을 제시하고 있다. 토픽 모델링을 통한 9개 주요 토픽에서 각각의 연구결과를 종합하여

살펴보면 다음과 같다:

첫째, 정밀의료 분야에서의 혁신과 기술 중심의 일자리 창출은 미래 지향적 연구의 중요한 주제로 부각된다. 인공지능, 스마트시티 등의 첨단 기술은 정밀의료 분야에서 중요한 역할을 수행하며, 이는 경제적 성장과 함께 일자리 창출에 기여할 것으로 예측된다. 둘째, 정밀의료 서비스 혁신과 빅데이터 활용은 환자 중심의 의료 서비스를 개선하고 효율성을 증진시킬 것으로 기대된다. 빅데이터 기술은 의료 데이터의 증가로 인한 관리의 어려움을 극복하며, 개인 맞춤형 의료 서비스의 확대 가능성을 보여준다. 셋째, 연구와 혁신은 정밀 의료 분야에서 핵심적인 역할을 수행하며, 새로운 치료법 및 의료 기기의 개발을 촉진한다. 다양한 분야 간의 연구 협력과 협업을 통해 효과적인 연구 생태계를 조성하는 것이 중요하다. 넷째, 빅데이터와 클라우드 기술을 활용한 정밀의료 데이터 관리는 의료 데이터의 증가로 인한 관리의 중요성을 강조하며, 표준화된 데이터 관리와 협력을 통해 향상된 의료 서비스를 제공할 수 있다. 다섯째, 유전체와 생물학적 연구는 암세포의 정확한 진단과 예방에 기여하며, 맞춤형 정밀의료 서비스의 발전을 지원한다. 유전체 정보를 기반으로 한 진단과 치료 전략의 발전은 더 나은 의료 결과를 제공할 것으로 예측된다. 여섯째, 바이오헬스와 융복합 기술은 현대 정밀의료 분야에서 사업 모델과 기술 개발을 혁신시키고 있다. 지자체와의 협업을 강화하여 지역 바이오산업 생태계를 육성하고 지역 경제의 발전을 촉진하는 것이 중요하다. 일곱째, 글로벌 시장에서의 경쟁과 협력은 정밀의료 기술과 산업의 발전에 큰 영향을 미치고 있다. 국가 간의 연구 협력과 산업 협력은 미래 지향적인 발전 방향을 제시하는 중요한 측면이다. 여덟째, 병원 간의 협력과 데이터 기술은 의료 서비스를 개선하고 환자 중심의 접근을 강화하여 데이터 기반의 평가를 통해 의료 시스템을 향상시키는 것이 중요하다. 아홉째, 신약 개발과 맞춤형 치료는 정밀 의료 분야에서 중요한 위치를 차지하고 있다. 유전자와 환자 데이터를 기반으로 한 맞춤형 치료는 진료 효과를 향상시키고 환자의 삶의 질을 향상시킬 것으로 기대할 수 있다.

각 토픽의 연구결과를 종합적으로 고려해보면 정밀의료 분야에서의 연구동향, 발전방향, 그리고 향후 전망을 더 깊게 이해할 수 있기 때문에 다음과 같은 시사점을 제시할 수 있다. 정밀의료 분야는 다학제적인 접근이 더욱 강조되어야 하는데 의료, 기술, 빅데이터, 유전체 등의 다양한 분야 간의 통합적인 연구와 협력이 필요하다. 그리고 국제적 시각에서의 연구 동향을 반영하여 향후 국제 시장에서의 혁신과 경쟁력을 강화할 수 있는 전략을 탐색하며, 국제적 협력과 연구결과의 공유 확산이 필요하다. 연구와 혁신은 윤리적 쟁점을 고려하여 이루어져야 하는데, 특히 유전체 및 빅데이터 활용과 관련된 윤리적인 고려사항에 대한 논의와 가이드라인이 더욱 필요하다. 마지막으로 연구결과를 토대로 국내외적으로 적절한 제도 및 정책 협조가 요구되므로 기술 혁신과 정밀의료 분야의 발전을 지원할 수 있는 제도적 지원이 필요하며, 국가 간의 협력체계를 강화하는 방안을 모색해야 한다.

본 연구는 정밀의료 분야의 연구동향을 정량적인 분석을 통해 탐색하여 현대 정밀의료 환경에서의 지속적인 발전과 향상을 위한 중요한 지침으로 활용될 수 있을 것이다. 다만 연구 동향은 시간에 따라 변화하며 진화하는데, 본 연구는 시기별 연구동향을 파악하지 못했다. 정밀의료는 사회적인 배경과 과학기술의 발전에 따라 바라보는 시각이 변화할 수 있기 때문에 향후에는 연구동향이 어떻게 발전되고 있는지를 추적하여 연구의 흐름 변화를 살펴볼 필요가 있겠다.

5. 감사의 글

본 과제(결과물)는 2023년도 교육부의 재원으로 한국연구재단의 지원을 받아 수행된 지자체-대학 협력기반 지역혁신 사업의 결과입니다.(2021RIS-001(1345370811))

References

- [1] Jun-hyun Song, Il-Gon, Kim Sun-Ju Ahn, Trends in Precision Medical Data Standardization, TTA Journal, (2017), Vol.172, pp.86-91.
Available from: <https://scienceon.kisti.re.kr/commons/util/originalView.do?cn=JAKO201764656419314&oCn=JAKO201764656419314&dbt=JAKO&journal=NJOU00292067>
- [2] Hye-Sun Yoon, Precision Medicine in Regulatory Perspectives: Triggered by the American Precision Medicine Initiative, Hanyang Law Review, (2018), Vol.35, No.4, pp.55-91.
DOI: 10.18018/HYLR.2018.35.4.055
- [3] Je-su Shin, Sang-ho Bae, Nam-yong Lee, Jin-ho Park, An Empirical Study of Healthcare & Medicine Information Protection Model for Precision Medicine Initiative, Journal of The Korea Society of Information Technology Policy & Management, (2016), Vol.8, No.4, pp.209-217.
UCI: G704-SER000003986.2016.8.4.001
- [4] Kyu-pyo Kim, Applying Precision Medicine in Clinical Practice, The Korean Journal of Medicine, (2020), Vol.95, No.6, pp.382-386.
DOI: 10.3904/kjm.2020.95.6.382
- [5] Se-young Moon, Gi-jeong Jang, Han-hae Kim, A Strategy of Success in Precision Medicine, KISTEP, (2016)
Available from: http://www.kistep.re.kr/c3/sub3_2.jsp?brdType=R&bbIdx=10502
- [6] Gyu-pyo Kim, Applying Precision Medicine in Clinical Practice, The Korean Journal of Medicine, (2020), Vol.95, No.6, pp.382-386.
DOI: 10.3904/kjm.2020.95.6.382
- [7] D. M. Blei, Probabilistic Topic Models, Communications of the ACM, (2012), Vol.55, No.4, pp.77-84.
DOI: 10.1145/2133806.2133826
- [8] A. P. Shiryayev, A. V. Dorofeev, A. R. Fedorov, L. G. Gagarina, V. V. Zaycev, LDA Models for Finding Trends in Technical Knowledge Domain, In 2017 IEEE Conference of Russian Young Researchers in Electrical and Electronic Engineering (EIConRus), (2017)
DOI: 10.1109/EIConRus.2017.7910614
- [9] L. Sun, Y. Yin, Discovering Themes and Trends in Transportation Research using Topic Modeling, Transportation Research Part C: Emerging Technologies, (2017), Vol.77, pp.49-66.
DOI: 10.1016/j.trc.2017.01.013
- [10] L. Liu, L. Tang, W. Dong, S. Yao, W. Zhou, An Overview of Topic Modeling and Its Current Applications in Bioinformatics, SpringerPlus, (2016), Vol.5, No.1, pp.1-22.
DOI: 10.1186/s40064-016-3252-8
- [11] D. Hall, D. Jurafsky, C. D. Manning, Studying the History of Ideas using Topic Models, In Proceedings of the 2008 Conference on Empirical Methods in Natural Language Processing, Association for Computational Linguistics, (2008)
Available from: <https://aclanthology.org/D08-1038>
- [12] <https://www.bigkinds.or.kr>, Jun 30 (2023)
- [13] Yong-min Baek, Text-Mining using R, Hanulm, (2020)
- [14] Min Song, Text Mining, Chongram, (2017)

- [15] D. M. Blei, Y. N. Andrew, I. J. Michael, Latent Dirichlet Allocation, *Journal of Machine Learning Research*, (2003), Vol.3, pp.993-1022.
- [16] Young-woo Kim, *Do It! R Text Mining*, EasyPublishing, (2021)
- [17] T. L. Griffiths, M. Steyvers, Finding Scientific Topics, *Proceedings of the National academy of Sciences*, (2004), Vol.101, pp.5228-5235.
DOI: 10.1073/pnas.0307752101