

융합기술분야 공백기술 분석을 통한 추진기술 분야 발굴 : 바이오마커 기술을 중심으로

권기석*, 박소연**, 안주명***, 서덕록****

I. 서론

공백기술 대상 중점 추진 분야 발굴 및 기술별 동향 조사를 통한 융합기술분야 신규 융합 R&D 사업 기획에 활용될 근거 도출하는 것이 중요한 과제로 인식되고 있다. 이에 본 연구는 '창조경제 실현을 위한 융합기술발전전략'상 15대 국가 전략 융합기술과 NTIS 과제 Mapping을 통한 공백기술 발굴에 초점을 두었다.

또한 R&D 투자확대, 과학기술의 첨단·융복합화 등의 연구개발 추세를 반영하여 증거기반의 보다 효과적인 전략적인 R&D 기획 요구, 미래부 중점 과학기술의 효과적 확보를 위해서 15대 국가 전략융합기술 추진 현황 분석을 통한 융합기술분야 공백기술 발굴이 필요하다.

따라서 본 연구에서는 국가전략융합기술의 중점기술별로 키워드를 도출하고 이를 바탕으로 국가적으로 어떤 세부 영역에 연구개발이 수행되고 있으며, 어떤 부분이 상대적으로 미흡한 지(공백기술)에 시계열 분석, 네트워크 분석 등을 수행하여, 향후 신규 융합R&D 기획에 활용하고자 한다.

연구의 구조와 흐름은 다음 그림과 같이 설명될 수 있다. 먼저, 융합기술데이터를 수집하였다. 국내의 경우 NTIS상에서 15대 국가전략 융합기술 중 3대 분야에 대한 과제정보를 수집했다. 이에 상응하는 미국의 연방정부 연구개발 투자 정보도 수집하였다. 이러한 데이터를 기반으로, 기술주제의 트렌드와 빈도 분석, 사회연결망 분석, 군집분석, 한국과 미국간 세부기술의 비교 분석을 실시하였다. 이러한 분석 결과를 바탕으로, 공백기술을 도출하고 신규 융합R&D 사업 기획의 시사점을 도출하였다.

II. 공백기술의 개념과 도출방법

공백기술의 개념과 도출에 있어, 공백기술(vacant technology)은 보통 해당 기술을 세부적으로 분류했을 때 이들 중 상대적으로 연구개발이 뒤쳐져 있는 세부분야 기술이라고 정의된

* 권기석, 한밭대학교 공공행정학과 교수, 042-821-1786, kiseok@hanbat.ac.kr

** 박소연, 한밭대학교 공공행정학과 학부생, 042-862-7887, calerpuck@hanmail.net

*** 안주명, 한국과학기술연구원(KIST) 박사후연구원, 02-958-4988, jmahn@kist.re.kr

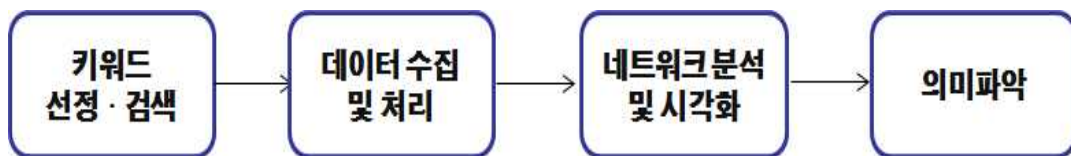
**** 서덕록, 한국과학기술연구원(KIST) 박사후연구원, 02-958-6922, dukrok@kist.re.kr

다(전성해, 2011; 박승, 2009; Lee et al., 2009). 그러나 이러한 ‘공백(vacancy)’을 어떤 기준으로 판단할 것인지에 대해서는 충분한 논의가 부족하다. 기존의 연구는 세부기술간 상대적 자원 투입이 부족한 분야로 가정하고 있으나, 자원 투입이 부족한지에 대한 판단 근거를 제시하기에는 문제가 있다고 본다. 따라서 본 연구에서는 동일 세부분야 선진국 기술과의 비교를 통해 ‘공백’에 대한 근거를 확보하고자 한다.

III. 방법론 및 데이터

1. 방법론

본 연구의 전체적인 분석 절차는 다음과 같다. 먼저, 중점기술의 키워드를 선정하고, NTIS에서 키워드 관련 데이터를 수집하고, 이의 서술통계, 네트워크 분석을 통한 시각화를 실시하고, 의미파악의 단계를 거쳐 시사점을 도출하고자 한다.



<그림1> 분석절차

단계별 분석방법은 다음과 같다.

[1단계 : 키워드 선정 및 데이터 수집] 이 단계에서는 중점 기술별 키워드 선정을 통해 NTIS에서 추출된 융합R&D 연구과제의 다양한 정보 (인력, 예산, 분야, 기간, 성과, 주제어)를 수집할 것이다. 또한, 기초적인 기술통계(descriptive statistics)분석이 수행될 것이다.

[2단계 : KrKwic을 이용한 빈출단어 추출] KrKwic을 활용한 빈도 분석을 통하여 중점 기술별 세부 연구주제를 도출할 것이다. KrKwic은 한국어 메시지 내용분석 프로그램으로서, 네덜란드 암스테르담 대학교의 Loet Leydesdorff 교수가 개발한 영어 메시지 분석 프로그램인 ‘Full Text’를 영남대학교 언론정보학과 박한우 교수가 한국어 분석을 위해 변형한 것이다. 각 문장들에 사용된 단어들을 개별 분류하여 각 단어의 출현빈도 및 단어들 간의 동시출현빈도 등을 분석하여, 언어 의미망 분석에 자주 활용되고 있다.

[3단계 : NetMiner를 이용한 네트워크 지도 작성 및 분석] NetMiner는 사회연결망 (소셜 네트워크) 분석 프로그램으로서 국내 IT 벤처기업인 (주)사이람이 개발한 사회연결망 분석 소프트웨어이다. 단어들 간의 관계를 나타내는 값에 따라 네트워크를 시각화하고 시각화 기능이 뛰어나고 중급 통계 분석 기능도 제공하여 네트워크 분석 분야에서 많이 사용되고 있다. 본 분석에서는 과제별 주제어를 바탕으로 네트워크 분석을 실시하여 15대 중점기술별로 어떤 연구과제가 진행되는지 파악할 것이다.

[4단계 : NodeXL을 이용한 군집화 분석] 네트워크분석 프로그램이기도 한 NodeXL은 화려한 군집화 시각화 기능을 가지고 있다. 네트워크 내에서 노드사이에 밀접한 관련성을 가지고 있는 경우, 하나의 군집으로 묶어주어 여러 군집이 출현할 수 있게 해준다. 이때 군집

의 형성에는 다양한 알고리즘이 적용될 수 있다.

[5단계 : 신규R&D 기획 시사점 도출 및 전문가 자문] 최종적으로 중점 기술별 융합R&D 연구개발 현황분석을 통해 향후 어떤 분야에 얼마나 연구자원을 투입할 것인지에 대한 시사점을 도출하고자 한다. 또한 정량적 분석 결과의 타당성을 높이기 위해 기술별 전문가 위원회를 구성하여 정성적 검토를 수행한다.

2. 데이터

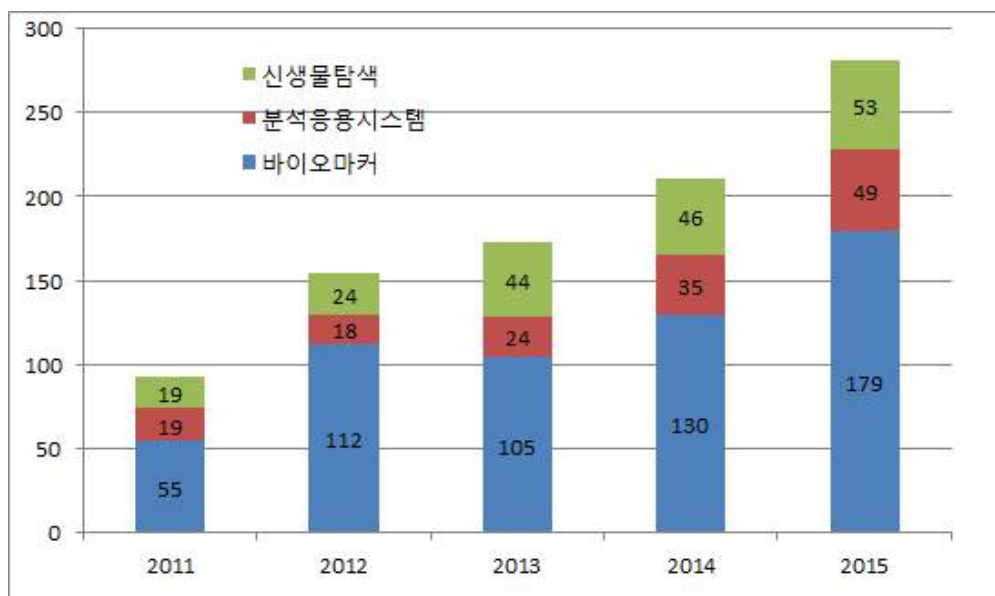
국가과학기술정보서비스(NTIS, <http://www.ntis.go.kr>)에서 유전체 관련 과제정보 3,701건. 이 중 검색어 색인을 통해 3대 중점기술 과제정보를 추출하였다.

미국연방정부R&D투자정보 (<https://www.starmetrics.nih.gov>) 데이터 검색을 통해, 유전체 관련 기술 중 가장 많은 빈도를 차지하고 기술적 정체성이 상대적으로 강한, 유전체 분야 중 바이오마커(7,326건)에 관련된 과제정보를 추출하였다.

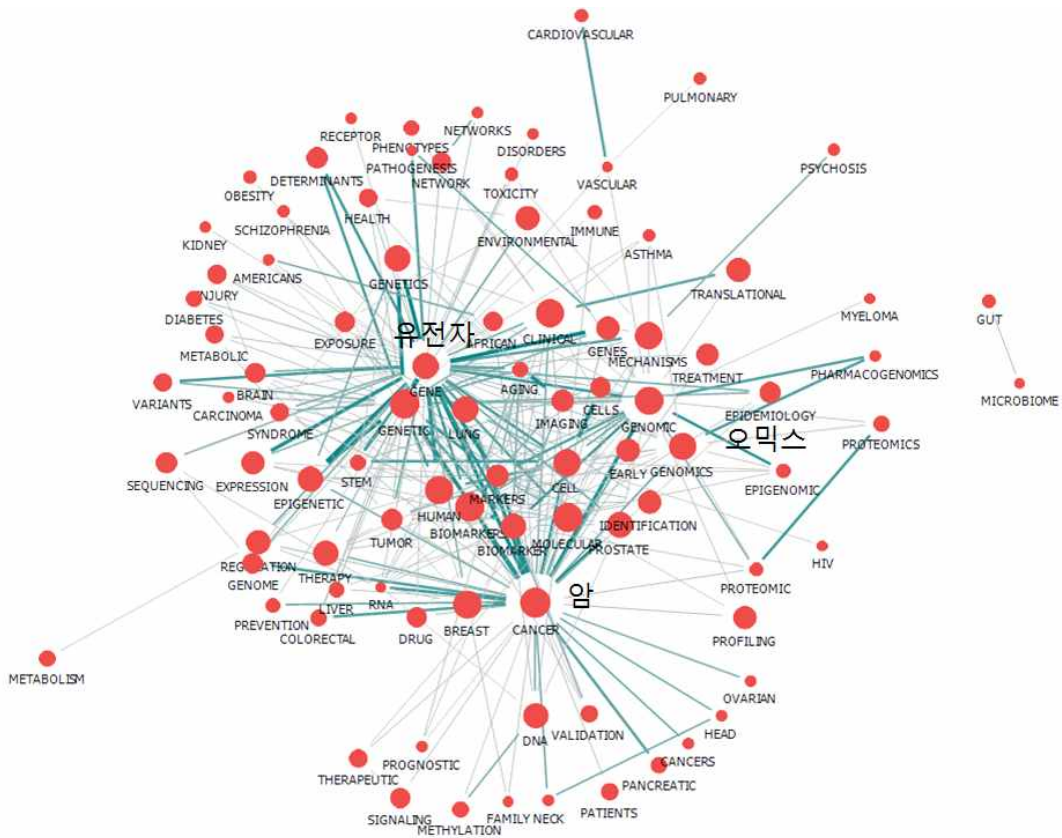
IV. 분석 결과

분석을 위해 국가과학기술정보서비스(NTIS)에서 2011년에서 2015년까지 5년간 수행된 유전체 관련 3,701건의 과제정보를 수집하였으며, ‘유전체 바이오마커 개발 기술(581건)’, ‘유전체 분석 응용시스템 개발 기술(145건)’, ‘유전체 기반 신생물 탐색 개발 기술(186건)’로 구성된 ‘유전체 정보 이용기술’은 912건으로 나타난다.

이러한 ‘유전체 정보 이용기술’ 관련 과제 수의 5년간 추이는 다음과 같음. 2011년 93개의 과제에서, 2015년 281개로 약 3배의 증가를 이룬 것을 알 수 있음. 세 핵심기술별 과제 수는 물론 유전체 관련 전체 과제 수가 꾸준히 증가한 것을 알 수 있음



<그림> 유전체 정보 이용기술의 과제 수 추이



<그림> 미국 과제 바이오마커 네트워크 (최근 5년간)

<표> 한국과 미국의 바이오마커 세부주제영역 비교

한국 주제어	공통 주제어 (좌상단은 한국 우세/우하단은 미국 우세)	미국 주제어
disease marker		GENOMICS
personalized		MECHANISMS
Pharmacogenomic		GENETICS
diagnosis		ENVIRONMENTAL
disorder		REGULATION
Medicine		EARLY
rare		EXPRESSION
medicine		GENES
screening		PROFILING
animal		IMAGING
Korea		MARKERS
metastasis		TREATMENT
microRNA		DETERMINANTS
risk		EPIDEMIOLOGY
agent		BRAIN
arthritis		SIGNALING
immunity		CELLS
mutations		EXPOSURE
chemotherapy		AFRICAN
(상위빈도 20위)		INJURY
		(상위빈도 20위)

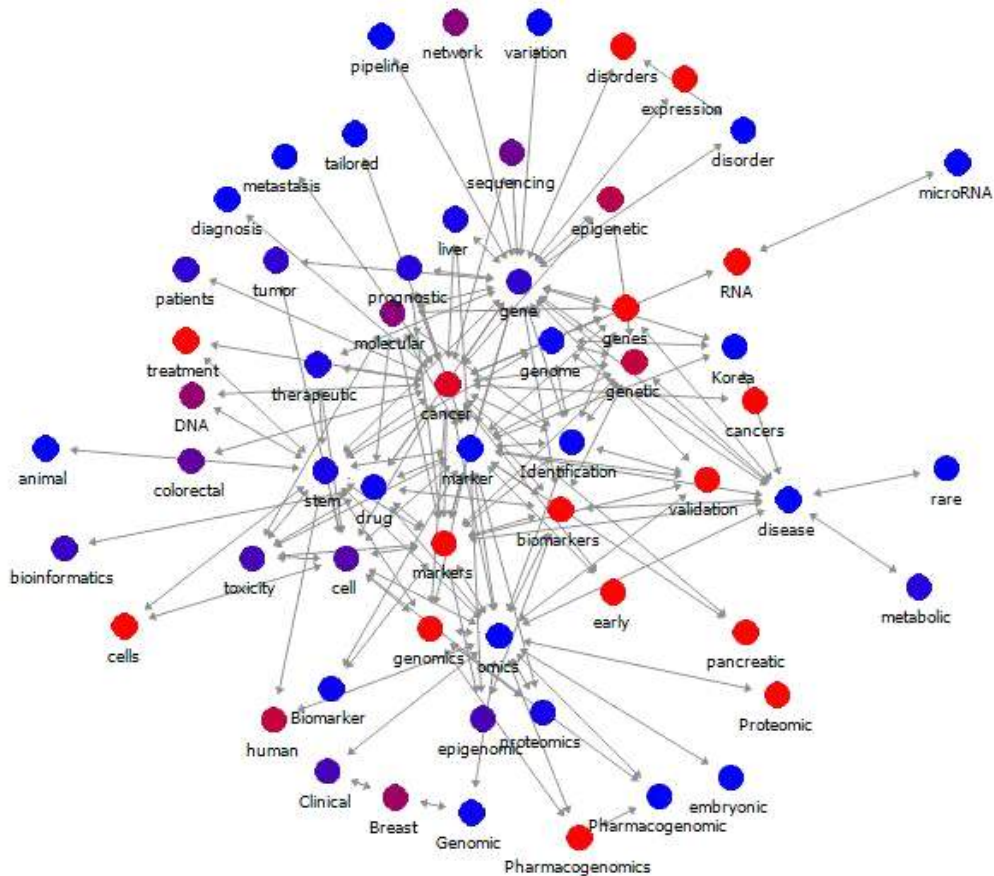
결론적으로 미국의 바이오마커 기술과 한국의 바이오마커 기술을 비교하였을 경우, 미국의 기술의 상대적으로 바이오마커의 기술이 여러 연령대와 신체의 다양한 부위를 대상으로 적용되고 있음을 알 수 있었다. 또한 암과 관련된 별도의 집중적 연구개발, 뇌질환 분야가 미국의 차별화된 점이라고 할 수 있다.

한국의 강점 연구영역 주제어로는 연구개발과 관련 주제어가 많이 분포하고 있으며, 현재 질환치료제에 많은 예산이 투입되고 있는 암 치료와 관련된 키워드(chemotherapy)가 질환 타겟으로 나타난다. 연구개발과 관련해서 관련된 유전체 염기서열 분석(NGS), 동물실험 관련 키워드(animal), 약물관련 키워드 (medicine, agent)가 존재하고 있다. 암 이외 질환 타겟으로 관절염(arthritis), 면역(immunity) 등이 존재함을 알 수 있다. 국가 9대과제 중에 하나인 정밀의료의 위한 맞춤형 약물개발 관련 키워드의 동의어가 많이 분포하고 있다. (personalized = individualized = tailored) 이런 것을 종합해 볼 때 한국의 강점 연구영역 주제는 아래와 같다고 볼 수 있다. personalized [individualized, tailored], animal, arthritis, immunity, chemotherapy, NGS, medicine [agent]

미국은 연구개발 방법론의 주제어보다는 질환 중심의 키워드(target-oriented)가 많이 분포하는 것이 특징이다. 질환도 대사성질환(metabolism, obesity), 뇌질환(brain, Alzheimer's, schizophrenia), 노화(aging), 심혈관 관련 질환(cardiovascular), 장 관련 질환(gut, microbiome), 암(cancer) 등으로 폭넓게 분포하고 있다. 또한 질환의 예방(prevention) 및 진단을 위한 분석기술(imaging) 관련 키워드로 존재한다. 이런 것을 종합해 볼 때 미국의 강점 연구영역 주제는 질환중심이며 다음과 같다고 볼 수 있다. imaging, brain, alzheimer's, metabolism, aging, prevention, cardiovascular, obesity, schizophrenia, HIV

공통 (경쟁 또는 협력) 연구영역 주제어로는 세계적으로 고령화 시대에 진입하게 되므로 이와 노화 관련 질환중심으로 협력이 좋은 모델로 될 것이다. 진단을 통한 질병의 예방, 개인맞춤형 약물은 앞으로 세계적인 트렌드가 되기 때문에 이 분야의 주제어가 중요하다. 이런 세계적 트렌드를 검토해보면, 1) 노화관련 질환 또는 대사성질환, 2) 예방, 3) 맞춤의학의 주제어가 협력분야의 주제어로 적합하다. 이를 다시 정리하면 다음과 같은 주제어로 표현할 수 있다. prevention, aging, metabolism, personalized = individualized = tailored

주제어 한미 공동 네트워크 분석을 표현한 다음 그림은 한국과 미국의 바이오마커 관련 주제어를 종합하여 네트워크를 구성한 것이다. 앞에서 논의한 바와 같이 한국은 다양한 방법론과 관련된 주제어가, 미국은 질환과 관련된 키워드가 주로 나타나는 것을 알 수 있다.



<그림> 바이오마커 주제어 한미 공동 네트워크

*붉은 색에 가까울 수록 미국 바이오마커의 주제어, 파란 색에 가까울 수록 한국 바이오마커 주제어를 나타냄

V. 요약 및 시사점

본 연구에서 제시한 방법론과 결과를 바탕으로 이와 같은 정책제언 및 본연구의 기여 부분을 제시할 수 있다. 본 연구는 융합기술의 공백기술 도출에 있어서, 연구 영역의 주제어 도출을 통한 네트워크분석, 시각화, 선진국과의 비교 분석 등을 결합하는 일련의 방법론으로 융합기술 뿐 만 아니라 타 연구 영역의 연구개발 기획에도 활용할 수 있을 것이다. 더 나아가 주제어를 기반으로 한 상대적 강점영역 (competitive advantaged area)의 도출과 이의 시각화는 향후 다양한 응용영역을 가지고 있다. 예를 들면, 국가별 강점연구 영역과 협력 가능 분야 도출이 가능하다는 점이다. 융합연구 영역에 도출에 있어서 정의된 주제어를 분야로 보고 학제간 연구(interdisciplinary research)를 주제어 겹침(keywords overlapping)의 정도에 따라 측정할 수 있다.

본 연구의 한계 및 향후 추가연구 방향을 다음과 같이 제시할 수 있다. 자료나 연구자원의 제한성으로 말미암아 하나의 세부분야에 대하여 연구를 진행하였으나, 향후 융합기술 전 분야로 확대해서 분석할 수 있을 것이다. 도출된 주제어는 물론 네트워크, 비교 분석에 있

어, 전문가의 질적인(qualitative) 판단 영역 내지 역할을 체계화하여 진행하여야 할 것이다. 연구결과 활용도를 높이기 위해서는, 분석결과 도출된 한국의 강점영역, 약점영역, 협력영역의 주제어와 가능 연구 분야에 대한 추가적인 연구개발 기획과 로드맵 및 국내외 협력전략이 제시되어야 한다.

참고문헌

- 전성혜(2011), 앙상블모형을 이용한 공백기술예측, 한국지능시스템학회 논문지, 21권 3호, pp.341-346.
- 박승(2009), 무인전투차량 기술개발 동향조사 및 분석, 한국산학기술학회논문지, 10권 7호, pp.1735-1739.
- 김요섭, 박상성, 장동식(2012), CLARA 알고리즘을 사용한 특허정보 분석: OLED기술 중심으로, 한국정보기술학회논문지, 10권 6호, pp.161~170.
- Lee, S., B. Yoon, Y. Park (2009), "An approach to discovering new technology opportunities: Keyword-based patent map approach," Technovation, vol. 29, pp. 481-497, 2009.

<웹사이트>

국가과학기술지식정보 <http://www.ntis.go.kr>

미국연방정부R&D투자정보 <https://www.starmetrics.nih.gov>