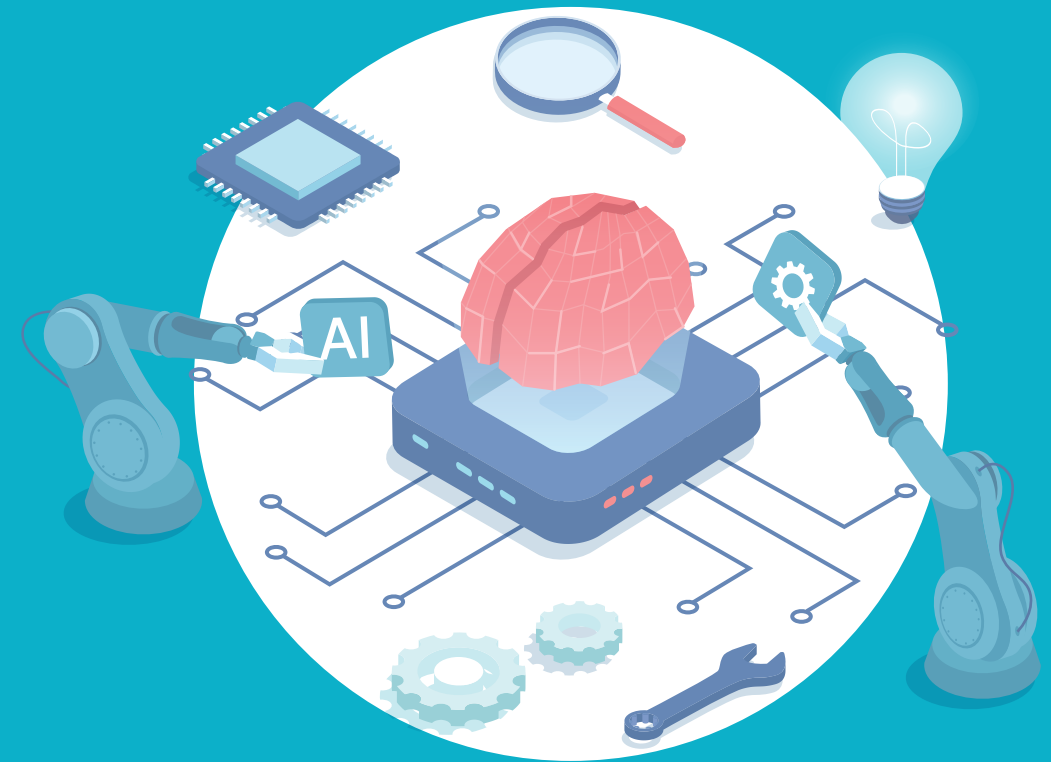




## 뇌연구 빅데이터 현황 및 연구 동향

손태권 · 구자욱 · 정민영 · 김기범 · 장원준 · 이종민



Big data may mean more information,  
but it also means more false information.

**빅데이터는 더 많은 정보를 뜻할 수 있지만 더 잘못된 정보를 뜻하기도 한다.**

- Nassim Taleb -



코로나 유행의 장기화로 힘들었던 우리는 이제 글로벌 금융위기에 비견될 정도의 심각한 경제 위기에 직면하고 있습니다. 지속된 경제적 불안정으로 사회적 우울을 초래하고, 많은 이들이 극단적 선택을 하는 등 현대 사회의 심각한 사회 문제로 대두되고 있습니다. 심지어 최근 안타까운 사고로 많은 이들에게 충격과 트라우마를 남기는 일도 있어,

국가 수준의 위기가 연속됨에 따라 우리나라 국민의 정신건강이 위협받고 있습니다.

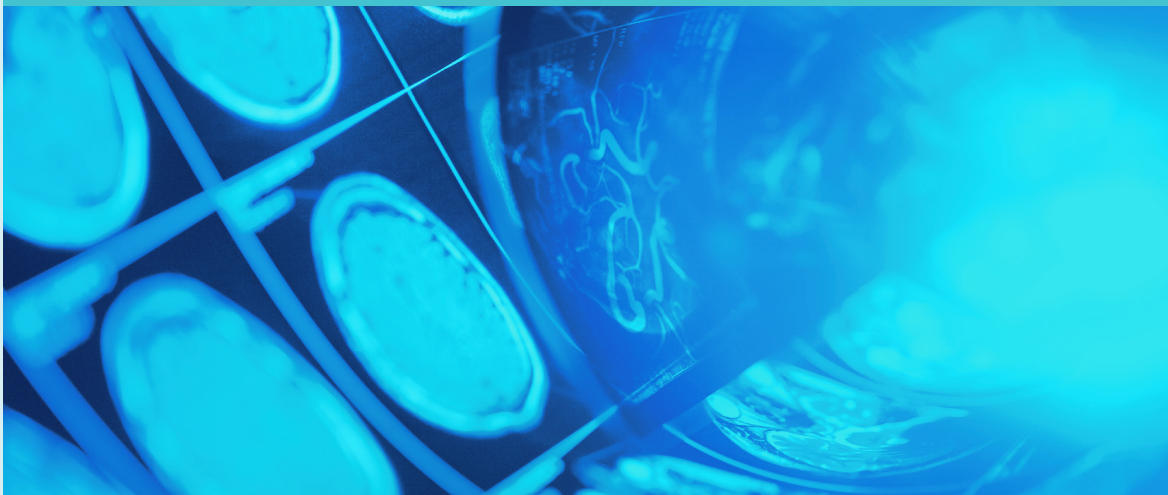
국민 정신건강뿐 아니라 각종 사회문제에 대응하는 뇌연구를 실천하기 위해, 많은 뇌과학 연구자들은 빅데이터를 활용하여 뇌를 이해하고자 다양한 시도를 하고 있습니다. 대규모 뇌영상 및 뇌연구 데이터를 활용해 치매, 우울증 등의 뇌질환들을 판별하기도 하며, 통증에 대한 객관적 지표를 마련하거나 발달 질환 및 희귀질환의 위험요소를 밝혀내기도 합니다. 이런 과정에서 데이터의 사소한 오류가 인류의 위대한 발견을 막기도 합니다. 이에 따라 가장 첫 시작점인 대규모 데이터를 표준화하고 좋은 품질의 데이터를 축적하는 것이 연구자들에게 아주 중요한 일로 인식되고 있습니다.

이에 한국뇌연구원은 제5호 Brain insight를 발간하여 뇌연구 데이터베이스와 관련한 현안에 대해 파악하고 우리가 앞으로 나아가야 할 방향에 대해 논의하는 장을 만들고자 합니다. ‘뇌연구자원 데이터베이스 동향’, ‘뇌신경정보 분석 플랫폼 동향’, ‘뇌연구자원 정보 연구 동향’에 대해 심도 있게 논의하고자 한국뇌연구원의 손태권, 구자욱, 정민영, 김기범 박사님과 한양대학교 이종민, 계명대학교 장원준 교수님이 집필에 참여하였습니다.

한국뇌연구원은 사회문제 및 미래 이슈 대응을 위한 뇌과학 기반 대응 방안을 마련하고자 꾸준히 노력하겠습니다. 아울러 국내 유일의 뇌연구 전문 국책연구기관으로서 국가와 사회에 기여하고 국민의 삶의 질 향상을 위해 임직원 모두 최선을 다하겠습니다. 감사합니다.

한국뇌연구원 원장 





## 목차/구성

### 5 제1장. 뇌연구자원 데이터베이스 동향

1. 자원-정보 통합형 데이터베이스
2. 뇌연구 데이터베이스
3. 주요 이슈 및 시사점

### 19 제2장. 뇌신경정보 분석 플랫폼 동향

1. 뇌연구 데이터 통합 분석 시스템
2. 국내외 뇌연구 데이터 플랫폼 개발 동향
3. 시각화 도구
4. 주요 이슈 및 시사점

### 29 제3장. 뇌연구자원 정보 연구 동향

1. 뇌영상/이미지 정보의 인공지능 분석
2. 오믹스 정보를 활용한 뇌질환 대상 신약 재창출
3. 뇌영상/이미지 정보와 오믹스 정보의 통합 분석
4. 주요 이슈 및 시사점





# 제1장

## 뇌연구자원 데이터베이스 동향

### 손태권

한국뇌연구원 한국뇌은행 /taekwon@kbri.re.kr

손태권 박사는 부산대 생물정보학 박사(2006)를 마치고, 국가생명연구자원 정보센터 연구원(2005-2007), BML의원 연구원(2012-2015), 서울대 종합약학연구소 책임연구원(2019-2020)을 거쳐 현재 한국뇌연구원 한국뇌은행 선임연구원(2020-)으로 재직중이다. 연구분야는 생물정보학, 메타분석, 시스템 생물학, 빅데이터통합, 정보의학 등이다.

### 구자욱

한국뇌연구원 글로벌 정서·중독 연구사업단 /jawook.koo@kbri.re.kr

구자욱 박사는 예일대 심리학과 행동신경과학 박사(2008)를 마치고, 마운트 시나이 아이칸 의과대학에서 박사 후 연구원(2008-2015)을 거쳐 현재 한국뇌연구원 글로벌 정서·중독 연구사업단 책임 연구원(2015-)으로 재직중이다. 연구분야는 감정, 인지, 정서 질환, 우울, 중독, 보상회로, 광유전학, 단일세포/세포특이적 전사체, 후생유전학, 생체 내 전기 생리학 등이다.

### 목차 및 구성

#### 요약

1. 자원-정보 통합형 데이터베이스
2. 뇌연구 데이터베이스
3. 주요 이슈 및 시사점

## 요약

뇌연구자원 데이터베이스는 뇌연구와 관련된 다양한 정보들을 집적, 통합하여 하나의 데이터베이스로 구축한 것으로, 기존의 생물정보 데이터베이스보다 넓은 범위의 데이터를 포함하고 있다. ‘뇌연구자원’은 생체 자체와 생체에서 유래된 뇌조직, 뇌척수액, 혈액 등을 포괄한 뇌연구소재와 이를 이용한 연구를 통해 얻은 정보인 뇌연구 데이터를 모두 포함하는 개념이다. 세부적으로, ‘뇌연구자원 데이터’는 뇌연구소재에 대한 연구를 통해 얻은 데이터와 이를 2차 가공한 데이터의 총칭으로 이미지/영상/신호 데이터, 오믹스 데이터, 임상정보 등이 있다.

뇌연구자원 데이터베이스는 생물정보 데이터베이스의 분류 기준인 1차<sup>1)</sup>, 2차<sup>2)</sup> 데이터베이스를 포괄적으로 다루고 있는 경우가 대부분이며, 기존의 생명의료정보 데이터베이스에 포함되어 있는 경우도 많다. 본 보고서에서는 각 기관에서 보유하고 있는 뇌연구소재 정보와 뇌연구 정보를 모두 포함한 ‘자원-정보 통합형 데이터베이스’와 뇌연구 자원 정보만 다루고 있는 ‘뇌연구 데이터베이스’를 모두 다루고자하며, 각 데이터베이스의 특성과 데이터베이스별 보유 정보를 정리하였다.

주: 1) 1차 데이터베이스 - GenBank, GEO 등 연구자의 데이터 등록 신청(submission)과 관리자의 승인 과정을 통해 구축된 데이터베이스

2) 2차 데이터베이스 - PROSITE, OMIM 등 1차 데이터베이스 내의 데이터를 큐레이터가 가공, 편집하여 구축된 데이터베이스

## 1

## 자원-정보 통합형 데이터베이스

가. UK Biobank (<https://www.ukbiobank.ac.uk/>)

UK Biobank는 영국 정부 복지부 지원을 받아, 암, 심장 질환, 뇌졸중을 포함, 광범위하고 인체에 치명적인 질병의 예방, 진단 및 치료를 개선하기 위해 주요 보건 연구 기관이 보유한 50만명(40~69세)의 건강기록을 기반으로 유전체 및 영상 데이터를 확보 및 공개하여 지속적인 서비스를 제공하고 있다.

## 1) 데이터 열람 방법 및 절차

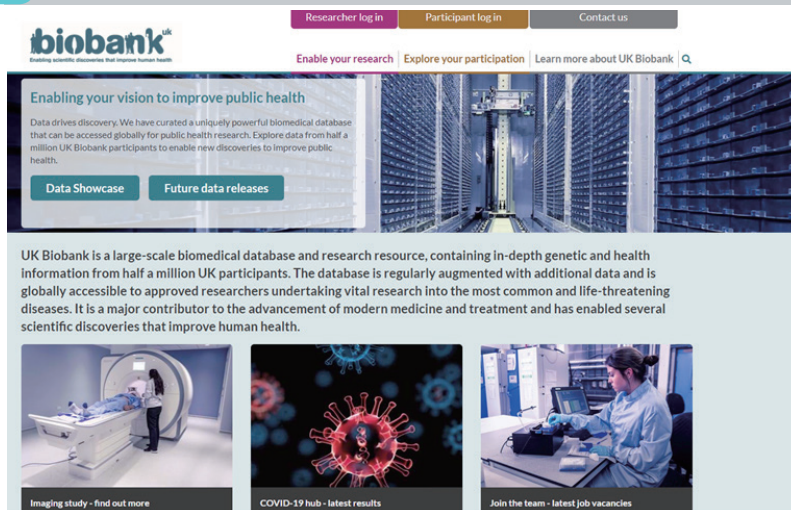
- (1) 데이터의 접근 방식: 다운로드(UK Biobank로부터 사전 동의 필요)
- (2) 공개 범위: 계정 생성 필요(데이터 사용 시 연간 업데이트 필요)
- (3) 계정 생성 절차: 등록-예비신청-주요신청-계약 총 14주 소요

표 1-1 계정생성절차 및 소요기간

생성 절차	소요기간
등록	신원 확인 2주
예비 신청서	검토 및 대응까지 4주
주요 신청서	검토 및 응답까지 4주
물질적 이전 계약	접속료 집행 및 처리까지 4주

## 2) 정보 플랫폼(UKB Managed Informatics Platform) 구성 및 데이터 서비스 종류

그림 1-1 UKB Managed Informatics Platform



### (1) 데이터 분량 및 분석 서비스

- 바이오데이터 분량을 위한 비식별처리
- 플랫폼 API를 통한 사용자 접근 어플리케이션 제공
- 유전체 데이터 분석을 위한 표현형 데이터 결과 제공
- 질병에 따른 표현형 데이터 분석 결과 제공

### (2) 데이터 보관 서비스

- 건강기록 및 질병 분류된 바이오 데이터를 저장 및 보관 서비스 제공
- 추가적으로 데이터를 업로드 및 다운로드 할 수 있도록 데이터 관리 서비스 제공

### (3) 분석을 위한 연산 인프라 제공 서비스

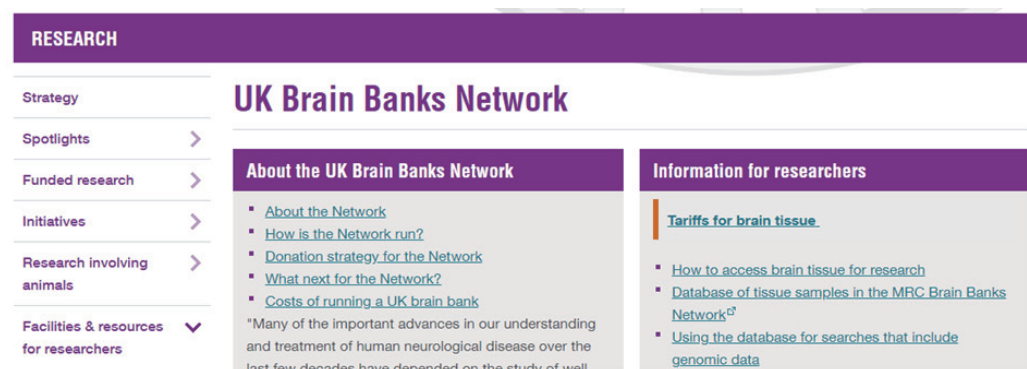
- 윈도우 및 리눅스 버전의 분석 연산을 위한 전산자원(CPU, GPU, memory) 등 인프라 제공 서비스

## 나. UK Brain Banks Network (UKBBN, <https://brainbanknetwork.ac.uk>)

영국 의학연구협회(MRC, Medical Research Council)가 이끄는 단체로 연구자들이 사용할 수 있도록 영국 뇌조직 자원에 대한 국가 네트워크를 구축하였다. 소속 뇌은행들은 질병 진단과 연구를 위해 일반인들이 기증한 사후 뇌와 중추신경계(CNS) 조직을 보관하고 연구 현황을 관리하고 있다.

그림 1-2

UK Brain Banks Network의 웹사이트



### 1) 데이터베이스 현황

UKBBN Database System은 UK 치매 플랫폼으로부터 치매 연구 데이터 세트를 확보하였고, 케임브리지 대학 부설 연구원에 의해 MRC 지원 연구 및 협력 연구에서 2,000 개 이상의 뇌 유전체 통합 DB를 구축하였다. UKBBN Database System은 10개의 영국 뇌은행의 전국 네트워크로, 표준화된 17,000개 이상의 두뇌 정보 데이터를 보유하고 있으며, 그 중 11,000개 이상의 뇌정보를 전 세계 학계 및 산업계를 대상으로 서비스를 제공하고 있다.

### 2) 검색 시스템 및 서비스 현황

UKBBN은 영국과 해외의 학술 및 산업 연구자들에게 조직 샘플을 공급하며, 네트워크의 모든 뇌은행에서는 연구 프로젝트 및 시험 연구에 조직 샘플을 제공하기전 데이터 공개를 위한 사전 승인을 받고 있다. 또한, 연구원 네트워크에서 조직 샘플의 세부 사항에 접근할 수 있도록 웹 기반 인터페이스 및 데이터베이스 접근 서비스를 제공한다.

연구자들은 치매, 파킨슨병, 자폐증 등 다양한 뇌질환 분야를 전문으로 하는 모든 은행의 조직을 통합·검색하여 데이터를 이용할 수 있다. 통합 DB system은 임상 진단, 병리학 결과, 인구 통계 데이터, 사후 검체 확보 시간 등 여러 검색 매개 변수의 조합을 통해 강력한 검색기능을 제공한다. 또한, 유전체 데이터와 임상 및 병리학적 결과를 포함하여 검색할 수 있고, 조직 샘플을 유전적 위험에 기초하여 선택할 수 있는 검색 기능도 제공한다.

그림 1-3

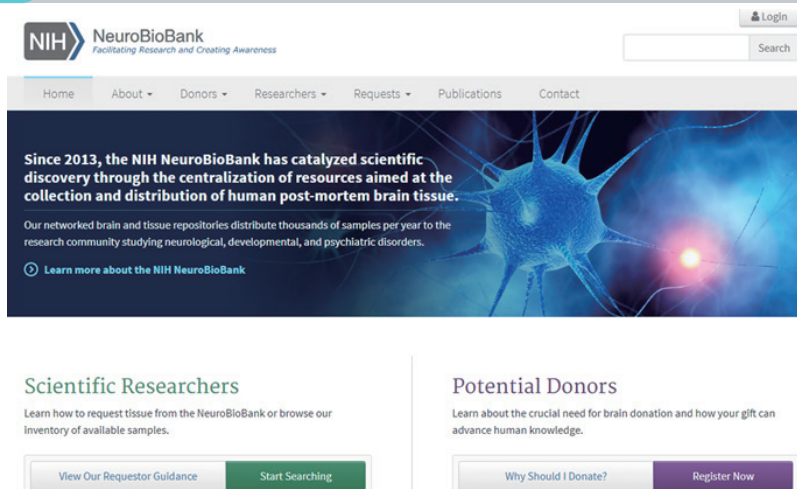
UKBBN 통합 DB 시스템

## 다. NIH NeuroBioBank(NIH NBB)

NIH가 2013년 9월에 설립한 NeuroBioBank는 사후 뇌조직 및 관련 생체 조직 검사를 신경계 연구에 활용하고자하는 연구자들을 지원하는 국가 기관이다. 신경 질환 및 뇌질환 장애에 이르는 광범위한 연구 분야에 고품질 사후 인간 뇌조직의 가용성을 높이고, 기증 자금을 활용한 규모의 연구 활성화를 목적으로 운영되고 있다. 인간의 뇌에 대한 연구뿐만 아니라 개별 모델 생물의 연구를 통한 건강 및 질병 관련 연구를 진행하고 있다. 특히, 신경학, 신경정신과, 신경 발달 질환, 장애를 포함한 검체를 확보한 NIH NBB는 6개의 생물 시료 저장 네트워크를 구성하여, 세계적 수준의 검체 수집 관리에 대한 중심 기관 역할을 하고 있다.

그림 1-4

NIH NeuroBioBank의 웹사이트

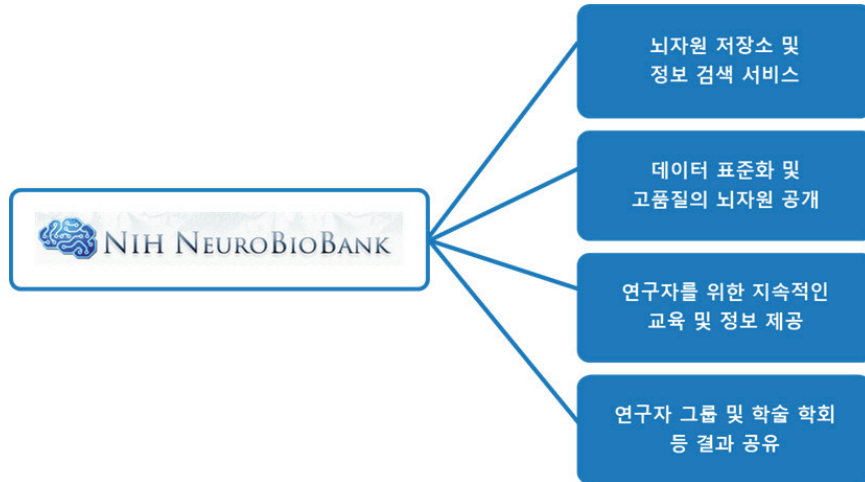


## 1) 데이터 서비스 종류

주요 서비스로 뇌자원의 저장 및 정보 검색, 데이터 품질관리(QC, Quality control) 및 표준화 관리, 교육 서비스, 학술활동 지원 등을 진행하고 있다.

그림 1-5

NIH NeuroBioBank의 주요 서비스

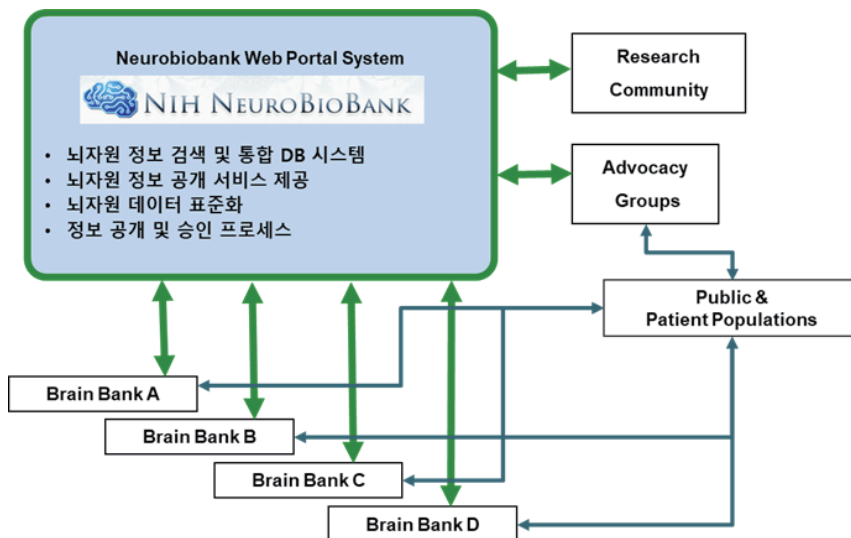


## 2) NIH NBB 데이터 포털 시스템

NIH NeuroBioBank의 포털 사이트는 미국 내 산재해있는 여러 생물자원 은행의 통합 사이트로서, 뇌자원 정보 통합 검색 시스템, 통합 데이터베이스, 데이터 표준화 정보, 정보 분양 절차 등을 제공하고 있다.

그림 1-6

NIH NeuroBioBank 포털 사이트의 구조와 주요 서비스



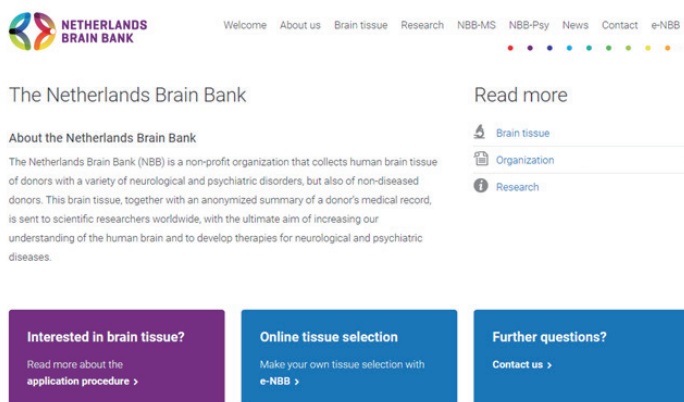


## 라. 네덜란드 뇌은행 (Netherlands Brain Bank, NBB, <https://www.brainbank.nl>)

네덜란드 뇌은행(NBB)은 다양한 신경 및 정신 질환을 가진 기증자의 인간 뇌 조직을 수집하는 비영리 단체로서, 개방형 액세스 정책을 준수하며 전 세계의 연구자를 대상으로 검체를 공개 분양하고 있다. NBB는 19개의 유럽 브레인 은행으로 구성된 BNE(Brain Net Europe) 컨소시엄(브레인 넷 유럽 II)의 데이터를 포함하여 운영되고 있다. 특히, NBB는 다양한 신경 및 정신 질환 증례에 해당하는 사망자의 부검샘플을 4,000개 이상 보유하고 있다. 그중, 사후 MRI 스캐닝 데이터는 다발성 경화증 (multiple sclerosis, MS) 환자의 사망 부검시 확보한 1cm의 관상 단층 이미지, 병변(lesion)의 MRI 영상(T1, T2, DI, FLAIR) 데이터 및 MRI 스캔 데이터를 보유하고 있으며 주로 조직학적 검증을 위해 활용되고 있다.

그림 1-7

네덜란드 뇌은행의 웹 사이트

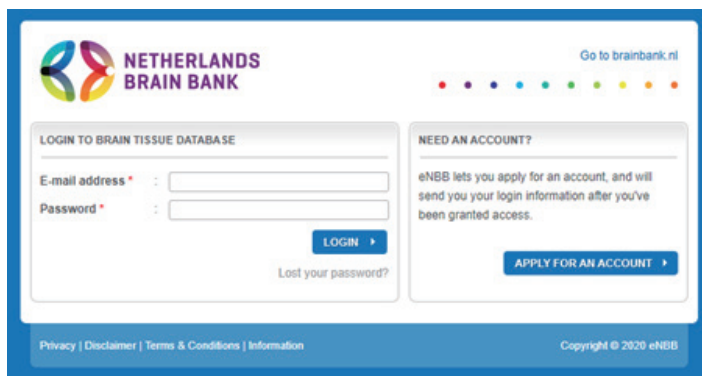


### 1) NBB Online Database System : e-NBB

e-NBB는 NBB 데이터베이스에 대한 온라인 액세스를 제공한다. 사용자는 현재 분양 가능한 뇌자원에 대해 다양한 데이터를 조회할 수 있다. 신청 가능한 샘플 유형은 피질(전두엽, 측두엽, 두정엽 등), 시상, 해마, 소뇌, 척수, 뇌척수액, 플라즈마, DNA 이다.

그림 1-8

e-NBB 데이터베이스





## 마. 한국뇌은행 네트워크 (<https://kbbn.kbri.re.kr>)

한국뇌연구원은 2014년에 과기정통부의 지원으로 한국뇌은행을 개소하였고, 2016년부터 7개소의 협력병원 뇌은행과 한국뇌은행네트워크를 출범하였다. 한국뇌연구원은 2016년~2017년까지 뇌연구자원 관리를 체계적으로 하기 위한 뇌연구자원 데이터 통합 DB인 ‘뇌연구자원 기록관리시스템(Brain Resource Archive Management System, BRAMS)’과 한국뇌은행네트워크 포털을 구축하여 2018년 6월부터 뇌연구자원 온라인 분양서비스를 운영하고 있다.

그림 1-9

한국뇌은행 네트워크의 뇌연구자원 검색시스템

뇌자원

**뇌자원 검색**

분양 기록 일자 안내

분양데스크

기탁데스크

### 뇌자원 검색

**검색조건**

구분	<input checked="" type="radio"/> 사신(뇌)자원 <input type="radio"/> 인체(뇌)자원		부검진단명		부검진단명 또는 KBBN Code를 입력하세요.	
자원구분	<input type="text" value="선택"/>	자원종류	<input type="text" value="선택"/>	보존온도	<input type="text" value="선택"/>	전뇌무게
성별	<input type="text" value="선택"/>	나이	<input type="text" value="선택"/> ~ <input type="text" value="선택"/>	장기	<input type="text" value="선택"/>	부위
DNA	<input type="text" value="선택"/> ~ <input type="text" value="선택"/>	RNA	<input type="text" value="선택"/> ~ <input type="text" value="선택"/>	RIN	<input type="text" value="선택"/> ~ <input type="text" value="선택"/>	
Thal phase	<input type="text" value="선택"/> ~ <input type="text" value="선택"/>	Broak stage	<input type="text" value="선택"/> ~ <input type="text" value="선택"/>	CERAD	<input type="text" value="선택"/> ~ <input type="text" value="선택"/>	Alzheimer's Disease Neuropathologic Change

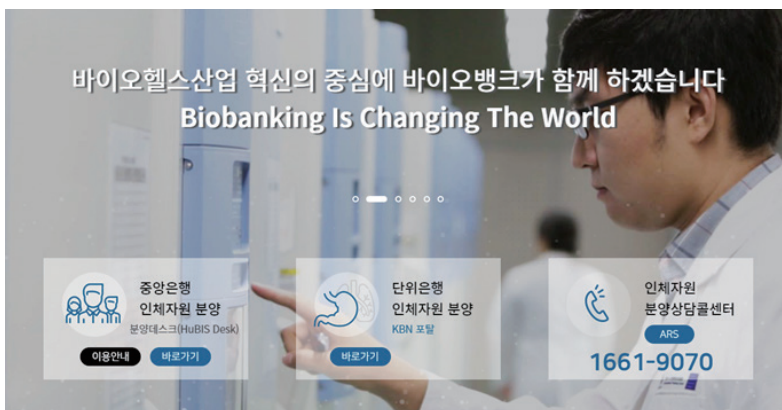
**검색결과(총 11,639건)**

	KBBN Code	임상진단명				부검진단명			
		자원구분	자원종류	보존온도	장기명	부위	방향	성별(나이)	전뇌무게
		사후경과시간	RNA	RIN	DNA	Broak stage	Thal phase	CERAD	Alzheimer's disease neuropathologic change
<input type="checkbox"/>		Normal brain with clinically interstitial lung disease							
<input type="checkbox"/>	KBBN01000553	조직	신선동결조직	-80℃	뇌	-	좌	여 (60-69)	1010g
		16h							

## 바. 국립중앙인체자원은행 (<https://nih.go.kr/biobank>)

국립중앙인체자원은행은 질병관리청 국립보건연구원에서 직접 운영하고 있는 인체자원은행으로 중앙은행 운영기획팀, 인체유래물 관리팀, 인체유래물 분양지원팀, 분양사무국, KBN(한국인체자원은행네트워크) 운영팀의 5개 부서로 구성되어있다. 국립중앙인체자원은행은 대학병원을 중심으로 인체자원단위은행을 지정하여, 각 인체자원단위은행에 예산을 지원함으로써 체계적인 질병 기반 인체자원 확보가 가능하도록 하였다. 현재 질병관리청 국립중앙인체자원은행(운영부서: 바이오뱅크과)을 중심으로 17개 대학병원소재 인체자원단위은행이 한국인체자원은행네트워크(Korea Biobank Network, KBN)를 구성하였으며, 대규모 인구집단 기반(중앙은행) 및 질병 기반 인체자원(17개 단위은행)을 수집, 관리하여 국내 연구자들에게 분양서비스를 제공하고 있다.

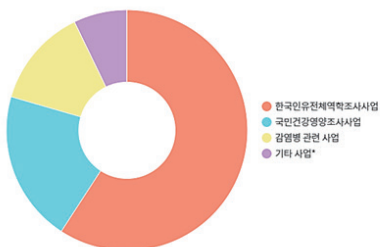
그림 1-10 국립중앙인체자원은행의 웹사이트



## 1) 데이터 현황

그림 1-11 국립중앙인체자원은행의 추진사업별 인체 자원 현황

○ 사업별 인체자원 수집 현황



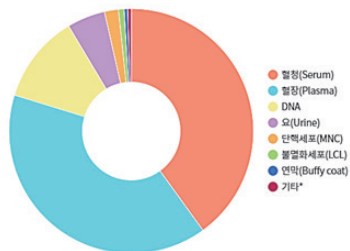
통계기준: 2019년 12월 31일 기준 누적치(As of Dec 31st 2019)

사업명	인원수
한국인유전체역학조사사업	246,966명
국민건강영양조사사업	84,891명
감염병 관련 사업	55,677명
기타 사업*	29,959명
총계	417,493명

\* 심뇌혈관 및 대사질환 원인연구, 데이터표준화에 기반한 오믹스 연구용 질환패널 자원 확보, 임신관련 합병증 예방을조사 및 위험인자 발굴 사업 등

그림 1-12 국립중앙인체자원은행의 인체유래물 종류별 자원 현황

○ 인체유래물 종류별 인체자원 수집 현황



통계기준: 2019년 12월 31일 기준 누적치(As of Dec 31st 2019)

유래물 종류	배달 수
혈청(Serum)	3,768,658
혈장(Plasma)	3,738,135
DNA	1,100,524
오(Urine)	473,899
단핵세포(MNC)	173,098
불활화세포(LCL)	71,312
연막(Buffy coat)	43,591
기타*	44,970
총계	9,414,187

\* 전혈, 적당, RNA 등

## 사. 인체자원 정보시스템(Human Biobank Information System, HuBIS)

인체자원 정보시스템은 인체자원단위은행 및 KBN 협력병원을 포함한 전국 57개 인체유래물 은행에서 공동 활용하는 시스템으로, 인체자원 정보시스템(HuBIS)은 기능별로 구분된 3개의 세부 프로그램으로 구성되어 있다. 각 프로그램은 사용자가 편리하게 이용하여 효율적인 업무가 가능하도록 매년 유지보수 사업을 진행하고, 사용자 지원을 위한 헬프데스크를 상시 운영하고 있다.

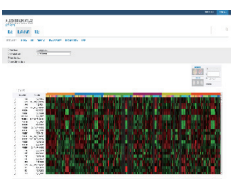
표 I-2 인체자원 정보시스템의 시스템 구조(단위시스템)

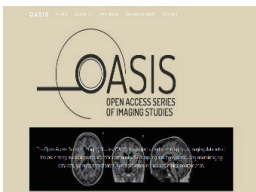

단위 시스템명	주요기능	사용자	비고
인체자원정보관리시스템 (HuBIS_Sam)	인체유래물의 수집, 보관, 분양까지의 전 과정에 대한 이력 관리	은행 자원관리자 및 정보관리자	국내 57개 인체유래물 은행에서 사용 중
원격지기탁지원시스템 (HuBIS_Tracker)	바코드라벨발행, 코호트사이트 관리	자원기탁자	국내 42개 기탁기관에서 사용 중
인체자원분양데스크 (HuBIS_Desk)	분양자원 검색, 온라인 상담, 분양 신청 및 심의, 자원활용성과 관리	연구자, 은행 분양담당자, 분양위원	국내 인체자원 연구자 1,000여명이 사용 중

## 2 뇌연구 데이터베이스

수십 종의 뇌연구 데이터베이스들 중 비교적 규모가 큰 데이터베이스들을 아래와 같이 정리하였다.

표 I-3 주요 뇌연구 데이터베이스(대형) 요약

데이터베이스	주요 내용
<p>Allen Brain Atlas</p> 	<ul style="list-style-type: none"> <li>• Allen Brain Atlas는 뇌의 광대한 연결정보를 단기간 내에 체계적으로 매핑하고 이 정보들을 사용자들에게 친숙한 방식으로 제공하여 누구나 쉽게 탐색할 수 있는 플랫폼을 구축, 2006년 9월부터 서비스 하고 있음</li> <li>• 생쥐 뇌의 모든 주요 구조에 대한 세포 단위 해상도 자료를 통해 생쥐의 약 22,000개의 유전자의 발현 정보를 담아 2006년 생쥐 뇌지도(Mouse Brain Atlas)를 작성하여 인터넷을 통해 공개하였다. 또한, 인간 뇌지도(Human Brain Atlas)도 완성하여 공개하였음</li> <li>• 웹에서 제공하는 데이터는 세포 유형 데이터베이스, 뇌 이미지 데이터베이스(Brain Observatory), 생쥐 뇌지도(Brain Atlas), 참조 뇌지도, 성장단계별 생쥐 뇌지도, 생쥐 척수 지도, 성인 및 성장단계별 NHP(영장류 모델) 지도, 인간 뇌지도, 성장 단계별 인간 뇌지도, 노화·치매 및 외상성 뇌손상(Traumatic brain injury, TBI), Ivy 아교모세포종 뇌지도 프로젝트 등으로 구분하여 제공하고 있음</li> <li>• <a href="https://www.brain-map.org">https://www.brain-map.org</a></li> </ul>

<p>Open Access Series of Imaging Studies (OASIS)</p> 	<ul style="list-style-type: none"> <li>• OASIS는 뇌의 신경 영상 데이터 세트를 과학계가 자유롭게 이용할 수 있도록 하는 프로젝트이며, 신경 영상데이터 세트를 컴파일하고 자유롭게 배포함으로써 기초 및 임상 신경 과학에서 미래의 발견을 촉진하는 것을 목적으로 개발되었음</li> <li>• <a href="https://www.oasis-brains.org">https://www.oasis-brains.org</a></li> </ul>
<p>BrainMaps</p> 	<ul style="list-style-type: none"> <li>• 2005년 5월에 출시된 BrainMaps.org는 영장류와 비영장류 뇌를 연속적으로 스캔한 이미지와 1억 4천만 픽셀 이상의 서브 마이크론 해상도 이미지 데이터를 보유하고 있으며, 사용자들은 다양한 검색 시스템을 활용하여 데이터를 열람할 수 있음</li> <li>• Macaca mulatta, Chlorocebus aethiops, Felis catus, Mus musculus, Rattus norvegicus, Tyto alba 등 다양한 생물종에 대한 완전한 뇌지도 데이터를 제공</li> <li>• 고해상도 뇌지도 데이터베이스로, 140 TB가 넘는 뇌 이미지 데이터를 온라인으로 제공하고 있으며, 뇌지도 데이터 탐색과 분석을 도와주는 프로그램들을 무료로 제공하고 있음</li> <li>• <a href="http://brainmaps.org">http://brainmaps.org</a></li> </ul>

그리고, 비교적 규모가 작은 데이터베이스들 중 많이 활용되고 있는 데이터베이스들을 아래 표에 요약하였다.

표 1-4 주요 뇌연구데이터베이스(중·소형) 요약				
데이터베이스	개체	데이터형식	질환 여부	제공되는 주요 정보
				웹사이트 주소
Alzheimer's Disease Neuroimaging Initiative	인간	MRI	정상, 알츠하이머	<ul style="list-style-type: none"> <li>• 구조적 MRI 이미지</li> <li>• <a href="http://adni.loni.usc.edu/">http://adni.loni.usc.edu/</a></li> </ul>
Brain Architecture Management System	생쥐(rat), 쥐(mouse), 인간	설명, 수치	정상	<ul style="list-style-type: none"> <li>• 신경 회로 정보</li> <li>• <a href="https://bams1.org/">https://bams1.org/</a></li> </ul>
Brain-Development	인간	MRI	태아, 신생아	<ul style="list-style-type: none"> <li>• 구조적 MRI 이미지 및 지도</li> <li>• <a href="http://brain-development.org/">http://brain-development.org/</a></li> </ul>
Braingraph	인간, 꼬마선충	그래프	정상	<ul style="list-style-type: none"> <li>• 인간 프로젝트 데이터로부터 계산된 뇌 그래프</li> <li>• <a href="http://braingraph.org/">http://braingraph.org/</a></li> </ul>
MSU Brain Biodiversity Bank	인간 외 62종	이미지	정상	<ul style="list-style-type: none"> <li>• 뇌지도, 뇌 및 MRI 이미지의 표준 섹션</li> <li>• <a href="http://www.brainmuseum.org/">http://www.brainmuseum.org/</a></li> </ul>
Brainomics Localizer Database	인간	이미지, 수치	정상	<ul style="list-style-type: none"> <li>• fMRI 및 MRI 스캔, 행동 데이터</li> <li>• <a href="https://neurovault.org/collections/4341/">https://neurovault.org/collections/4341/</a></li> </ul>

International Neuroimaging Data-sharing Initiative	인간	기능성 연결	정상, 질환	<ul style="list-style-type: none"> <li>• 여러 그룹의 기능적 연결 데이터</li> <li>• <a href="http://fcon_1000.projects.nitrc.org">http://fcon_1000.projects.nitrc.org</a></li> </ul>
Invertebrate Brain Platform	무척추 동물 47종	이미지	정상	<ul style="list-style-type: none"> <li>• 무척추 동물 신경계 해부 사진</li> <li>• <a href="https://invbrain.neuroinf.jp/">https://invbrain.neuroinf.jp/</a></li> </ul>
Mouse Brain Library	생쥐	이미지	정상	<ul style="list-style-type: none"> <li>• 뇌지도, 생쥐 뇌의 염색된 조직 절편</li> <li>• <a href="http://www.mbl.org/">http://www.mbl.org/</a></li> </ul>
Neurodata	생물종	이미지	다양	<ul style="list-style-type: none"> <li>• 대용량 데이터 세트, 아틀라스 및 연결체 연구</li> <li>• <a href="https://neurodata.io/">https://neurodata.io/</a></li> </ul>
Neuromorpho	인간, 생쥐, 쥐, 원숭이	이미지	정상	<ul style="list-style-type: none"> <li>• 실제 뉴런의 3D 모델</li> <li>• <a href="http://neuromorpho.org/">http://neuromorpho.org/</a></li> </ul>
Neuromorphometrics	인간	MRI	정상	<ul style="list-style-type: none"> <li>• 수동으로 라벨링 된 MRI 뇌 스캔</li> <li>• <a href="http://www.neuromorphometrics.com">http://www.neuromorphometrics.com</a></li> </ul>
NeuronDB	인간	설명	정상	<ul style="list-style-type: none"> <li>• 뉴런 속성 및 분류 데이터베이스</li> <li>• <a href="https://senselab.med.yale.edu/NeuronDB/">https://senselab.med.yale.edu/NeuronDB/</a></li> </ul>
Whole Brain Atlas	인간	이미지	정상, 질환	<ul style="list-style-type: none"> <li>• 뇌지도, 구조적 MRI 이미지 및 PET 이미지</li> <li>• <a href="http://www.med.harvard.edu/AANLIB/">http://www.med.harvard.edu/AANLIB/</a></li> </ul>

### 3 주요 이슈 및 시사점

#### 가. 데이터 표준화 현황

현재 국내에서는 한국인 표준 뇌지도, 뇌질환 진단 표준화 및 임상 진료지침 마련에 관한 연구를 진행하고 있으며, 향후 뇌연구 데이터에 대한 활용성 증대를 위해 표준화된 프로토콜이 필요하다는 인식이 높아지고 있다. 연구 과제를 통해 진행되고 있는 국내 뇌연구 데이터의 표준화를 위해서는 수집한 뇌연구 데이터의 표준화를 도출하기 위해서 데이터 원칙을 정립하고, 이에 따라서 표준체계를 설계하여 뇌연구 데이터의 표준화된 프로토콜을 수립하여야 한다.

범용적인 표준화를 위해서는 개념정의, 중복제거, 글로벌 표준준수 등의 원칙을 준수할 필요가 있다. 표준체계 설계에서는 분류체계 정립, 코드체계 및 명명규칙 정립, 속성항목 및 값 표준화, 메타데이터 및 데이터사전 구축 등을 수행하는 활동으로 확립된 표준화 프로토콜을 수립하여야 한다. 또한, 100개 이상의 해외 뇌연구 관련 데이터베이스를 수집하여 메타정보를 제공하는 NITRC(Neuroimaging Informatics Tools and Resources Clearinghouse)를 참조하여 뇌연구 데이터의 표준화 프로토콜에 반영할 필요가 있다.

뇌연구 데이터에 대한 일관된 데이터 형식 및 규칙을 적용하고 주기적인 품질 관리(QC)를 통해서 표준화에 대한 관리 프로세스 체계를 적용 하며, 뇌연구 데이터 표준을 지속적으로 유지

해서 데이터의 활용성을 향상시킬 수 있도록 해야 할 것이다. 뇌연구 데이터 표준화 프로세스는 표준화 체계 수립, 진단대상 정의, 품질측정, 품질측정 결과분석, 품질개선의 순서로 이루어지며 이 과정을 반복하며 지속적으로 표준화를 개선해나가야 한다.

## 나. 뇌연구 데이터의 활용 및 보안을 위한 정책 수립 및 데이터 거버넌스 규정

뇌연구 분야의 특성상 저장되는 데이터는 정보 제공자와 실험 방법에 따라 다양한 종류의 연구자원이 생산, 활용되므로, 경우에 따라서 임상·의료 정보에 대한 관리 체계가 필요할 수 있다. 임상·의료 데이터인 경우는 행정자치부의 개인정보 보호 가이드라인과 보건 복지부의 의료법 시행 규칙을 준수하면서 영상 데이터 관리 시스템 구성에 반영하여야 한다.

## 다. 개인정보보호 및 보안 기술

### 1) 위조 및 변조 방지 방안

접속기록에 대한 위·변조를 방지하기 위해 DVD-ROM, WORM(Write Once Read Many) Disk 등과 같은 덮어쓰기 방지 매체를 사용하는 것이 필요하다. 또한, 접속기록을 수정 가능한 매체(HDD 또는 테이프 등)에 백업하는 경우에는 무결성 보장을 위해 위·변조 여부를 확인할 수 있는 별도의 장비에 보관, 관리하는 것이 요구된다.

### 2) 뇌기증자의 임상정보 보존 및 보안 강화

의료법 시행규칙 제15조 제1항의 진료에 관한 기록은 마이크로필름이나 광디스크 등에 원본대로 수록하여 보존 가능하다. 보안을 강화하기 위해서는 임상·의료 데이터에 대한 WORM 스토리지와 같은 보안이 강화된 저장장치를 활용할 필요가 있다. 또한, 데이터의 위변조 방지에 대한 개인정보보호법 규정을 충족해야하며, 다양한 보안 위협, 삭제 및 훼손의 위협으로부터 데이터의 안전한 백업이 보장되어야 한다. 의료영상 저장전송시스템(Picture Archiving and Communication System, PACS) 영상 등 파일로 존재하는 다양한 데이터에 대한 안전한 백업이 보장되어야하며, 시스템 내부통제 및 보안사고 추적을 위한 통합 로그 관리가 요구된다

---

---

## 참고문헌

1. 김모모, 생명공학정책연구센터, “뇌과학연구 및 활용기술(BT 기술동향 보고서).” 총서 제189권. 2012.10.
2. 심원식, 성균관대학교. “국가 차원의 연구데이터 관리체계 구축을 위한 로드맵 제안 - 영국 사례 분석을 중심으로 -.” 한국문헌정보학회지 4(4): 355-378. 2015.
3. WIKIPEDIA, “List of neuroscience databases.” ([https://en.wikipedia.org/wiki/List\\_of\\_neuroscience\\_databases](https://en.wikipedia.org/wiki/List_of_neuroscience_databases))
4. Li, Xin, and Huadong Liang. “Project, toolkit, and database of neuroinformatics ecosystem: A summary of previous studies on "Frontiers in Neuroinformatics".” Frontiers in neuroinformatics vol. 16 902452. 26 Sep. 2022, doi:10.3389/fninf.2022.902452
5. Nayak, Losiana et al. “Computational neuroscience and neuroinformatics: Recent progress and resources.” Journal of biosciences vol. 43,5 (2018): 1037-1054.
6. Ryan, Jim, and Kaitlyn Casimo. “A Course-Based Research Experience Using the Allen Brain Map: From Research Question to Poster Session.” Journal of undergraduate neuroscience education : JUNE : a publication of FUN, Faculty for Undergraduate Neuroscience vol. 19,2 A260-A266. 20 Jun. 2021.
7. Sunkin, Susan M et al. “Allen Brain Atlas: an integrated spatio-temporal portal for exploring the central nervous system.” Nucleic acids research vol. 41,Database issue (2013): D996-D1008. doi:10.1093/nar/gks1042.
8. Sweeney, Elizabeth M et al. “OASIS is Automated Statistical Inference for Segmentation, with applications to multiple sclerosis lesion segmentation in MRI.” NeuroImage. Clinical vol. 2 402-13. 15 Mar. 2013, doi:10.1016/j.nicl.2013.03.002.
9. Mikula, Shawn et al. “BrainMaps.org - Interactive High-Resolution Digital Brain Atlases and Virtual Microscopy.” Brains, minds & media : journal of new media in neural and cognitive science and education vol. 3 (2008): bmm1426.
10. Mikula, Shawn et al. “Internet-enabled high-resolution brain mapping and virtual microscopy.” NeuroImage vol. 35,1 (2007): 9-15. doi:10.1016/j.neuroimage.2006.11.053.
11. Kennedy, David N et al. “The NITRC image repository.” NeuroImage vol. 124,Pt B (2016): 1069-1073. doi:10.1016/j.neuroimage.2015.05.074.
12. 의료기기산업팀, 한국보건산업진흥원, "인공지능(AI) 기반 의료기기 현황 및 이슈", 보건산업브리프 - 의료기기 주요 이슈 분석, 2018.

## 제2장 뇌신경정보 분석 플랫폼 동향

### 손태권

한국뇌연구원 한국뇌은행 /taekwon@kbri.re.kr

손태권 박사는 부산대 생물정보학 박사(2006)를 마치고, 국가생명연구자원 정보센터 연구원(2005-2007), BML의원 연구원(2012-2015), 서울대 종합약학연구소 책임연구원(2019-2020)을 거쳐 현재 한국뇌연구원 한국뇌은행 선임연구원(2020-)으로 재직중이다. 연구분야는 생물정보학, 메타분석, 시스템 생물학, 빅데이터통합, 정보의학 등이다.

### 정민영

한국뇌연구원 인지과학연구그룹 /minyoung@kbri.re.kr

정민영 박사는 일본 오사카대 아동 발달학 박사(2015)를 마치고, 일본 학술진흥회(JSPS) 리서치 펠로우(2014-2015), 하버드 의과대학 정신과 박사 후 연구원(2015-2017), 후쿠이 대학 의학부 신경정신과 조교수(2019-2021)를 거쳐, 현재 한국뇌연구원 인지과학연구그룹 선임연구원(2021-)으로 재직중이다. 연구분야는 사회인지, 감정, 아동발달, 자폐스펙트럼 장애, 주의력결핍 과다행동장애 뇌 발달 등이다.

### 목차 및 구성

#### 요약

1. 뇌연구 데이터 통합 분석 시스템
2. 국내외 뇌연구 데이터 플랫폼 개발 동향
3. 시각화 도구
4. 주요 이슈 및 시사점



## 요약

1991년 유럽은 '유럽의 뇌 연구 10년'(European Decade of the Brain)을 선언하고 뇌 연구를 본격 시작하였으며, 2013년 뇌를 슈퍼컴퓨터에서 시뮬레이션하는 휴먼 브레인 프로젝트(Human Brain Project, 2013년~2024년 / 약 1.4조원 투자)를 유럽 내 100여개 연구기관이 공동으로 진행하여 뇌 신경망 연구와 인공지능 플랫폼 개발을 추진하고 있다. Human Brain Project의 9가지 우선순위 연구영역은 뇌 탐색, 뇌 시뮬레이션, 가상 뇌, 인지 이해, 의학, 로봇, 대규모 컴퓨팅, 사회·윤리·철학적 지원 및 통합적 접근 분야이다. 또한, 미국 알렌 뇌과학연구소는 2003년도에 설립한 비영리 연구기관으로 인간의 뇌 작동 원리 이해 등 뇌과학의 중요 이슈를 발굴하여 글로벌 선도 연구를 수행하고 있으며, 오픈데이터 플랫폼 구축을 위해 5천억원 투입(2003년~2014년)하고, NIH는 뇌지도 작성을 위한 BRAIN Initiative 사업으로 연간 600억원 규모의 예산을 투입하고 있다. 현재, 인간과 마우스를 비롯한 다양한 생물종에 대한 뇌연구 데이터를 보유한 시스템이 다양하게 구축되어 서비스되고 있다. 주로 MRI, fMRI와 같은 뇌영상 데이터와 뇌신경 이미지 정보, 오믹스 정보의 열람과 분석 기능을 제공하고 있다. 본 장에서는 국내외 뇌연구 데이터 플랫폼 개발 동향에 대해 알아보고, 세계 각국의 다양한 분석 플랫폼 중 많이 이용되고 있는 BigBrain과 JuBrain에 관해 아래에 좀 더 자세히 설명하겠다.

## 1

## 뇌연구 데이터 통합 분석 시스템

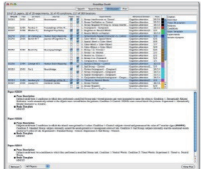
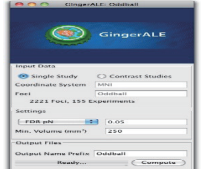
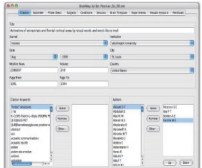
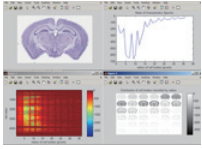
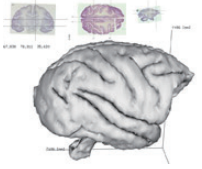
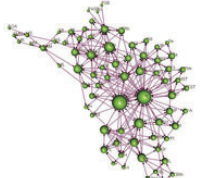
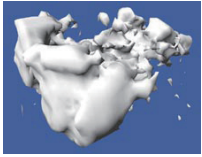
현재, 뇌신경정보 분석 플랫폼 구축은 미국과 유럽을 중심으로 활발하게 진행되고 있다. 유럽 휴먼 브레인 프로젝트(HBP)에서는 EU의 지원하에 2013년도에 시작된 10년간의 연구프로젝트로 스위스 로잔공대 주관으로 설치류 및 인간 뇌 모델구축을 목표로 다양한 IT 인프라의 개발을 진행하고 있다. HBP는 인간 뇌 이해의 가장 큰 장애물을 통합되지 않은 다양한 뇌 연구 데이터로 규정하고, 최신 IT 기술을 이용하여 다양한 뇌 연구 데이터를 통합하고 하나의 플랫폼에서 뇌를 모델링하고자 한다. HBP의 세부 목표는 첫째, 유전체로부터 행동에 이르기까지 모든 단계의 생물학적 뇌조직에 대한 지식과 정보를 통합하여 인간 뇌의 다층적 구조를 이해하는 것(future neuroscience)이다. 둘째, 뇌신경 회로와 컴퓨터 원리를 기반으로 신경모방기술, 신경로봇기술을 발전시키고, 뇌 시뮬레이션, 로봇, 자율신경 제어 등을 위한 슈퍼 컴퓨팅 기술을 개발하는 것(Future Computing)이다. 셋째, 다층적 임상 데이터 기반의 신경질환, 정신질환 지도를 개발하고, 이를 활용하여 뇌 질환을 진단하고 분류하는 것이다. 세부 프로젝트로는 신경정보학, 의료 정보학, 뇌 시뮬레이션, 초고성능 컴퓨팅, 뉴로모픽 컴퓨팅(neuromorphic computing), 신경 로봇학(neurobotics) 등 13개 분야에서의 ICT 플랫폼 구축이다.

미국의 앨런 연구소(Allen Institute)는 미국국립보건원(National Institutes of Health)의 지원으로 인간 두뇌의 세포 지도를 만드는 프로젝트를 주도하고 있으며, 뇌 세포의 위치, 특징 및 기능을 매핑하고있다. 살아있는 뇌 조직에서 전기생리학, 해부학, 유전체 등 다양한 뇌연구 데이터를 통합한다. 이 프로젝트는 4개의 세부 프로젝트로 진행되고 있다. 첫째, 생쥐의 발달 단계별로 두뇌를 매핑한다. 둘째, 동물이 시각적 환경을 인식하고 이에 반응할 때 뉴런의 활동을 포착하여, 시각 정보를 처리하는 마우스 뇌 영역과 뇌 세포 유형, 기능을 연결한다. 셋째, 공개 웹 기반 지식 플랫폼을 구축하고, 가시화 시스템을 개발, 운용하는 것이다. 넷째, NIH의 뇌 지도 작성 연구를 지원하기 위한 센터를 구축하는 것이다.

본 보고서에서는 뇌연구 데이터를 통합 분석할 수 있는 시스템들 중 비교적 많이 활용되고 있는 브레인맵(BrainMap)<sup>1)</sup>과 브레인맵스(BrainMaps)<sup>2)</sup>의 소프트웨어들을 중심으로 아래와 같이 정리하였다.

주: 1) 텍사스 대학교 헬스 사이언스 센터(San Antonio)의 Research Imaging Institute에서 운영

2) 캘리포니아 대학교 데이비스 캠퍼스에서 운영

시스템	특징	화면
Sleuth	관심 있는 뇌관련 논문을 검색해주는 프로그램 메타 데이터를 읽어 좌표로 시각화하는 기능 제공 프로그램 주소 : <a href="http://brainmap.org/sleuth/">http://brainmap.org/sleuth/</a>	
GingerALE	ALE 메타 분석을 수행하는 프로그램 Talairach 또는 MNI 공간 좌표에 GingerALE을 사용하여 변환가능하고 icbm2tal 변환을 사용하여 MNI와 Talairach 공간 간의 좌표 조정 기능을 제공 프로그램 주소 : <a href="https://brainmap.org/ale/">https://brainmap.org/ale/</a>	
Scribe	좌표 및 메타 데이터를 추출하는 프로그램 논문으로 발표된 기능성 신경 영상 데이터를 보유 프로그램 주소 : <a href="https://brainmap.org/scribe/">https://brainmap.org/scribe/</a>	
BrainMaps Analyze	BrainMaps.org의 이미지 데이터를 분석해주는 프로그램 다양한 형태의 그래프 시각화를 지원 프로그램 주소 : <a href="http://brainmaps.org/index.php?p=desktop-apps-brainmaps-analyze">http://brainmaps.org/index.php?p=desktop-apps-brainmaps-analyze</a>	
BrainMaps B3D	시각화된 뇌의 3D 구조 위에서 고해상도의 2D 이미지 데이터를 확인할 수 있는 웹 프로그램 프로그램 주소 : <a href="http://brainmaps.org/index.php?p=desktop-apps-b3d">http://brainmaps.org/index.php?p=desktop-apps-b3d</a>	
Nodes3D	3D 그래프 시각화 프로그램 프로그램 주소 : <a href="http://brainmaps.org/index.php?p=desktop-apps-nodes3d">http://brainmaps.org/index.php?p=desktop-apps-nodes3d</a>	
Make Isosurf Volume	뇌 이미지 세트로부터 볼륨파일을 생성해주는 프로그램으로, 생성된 파일은 표면 삼각측량 프로그램인 IsoSurf에서 사용 가능 프로그램 주소 : <a href="http://brainmaps.org/index.php?p=desktop-apps-make-isosurf-volume">http://brainmaps.org/index.php?p=desktop-apps-make-isosurf-volume</a>	

### 가. 심리-뇌영상 데이터를 활용한 정신건강 해외 데이터 공유 사례

현재 뇌영상 이미지를 통해 빅데이터와 개방 과학(open science) 정신에 입각해 데이터를 구축하고 공유를 주도하는 국가로는 미국, 캐나다, 영국과 아일랜드를 중심으로 한 EU, 중국, 일본이 있으며, 국가별 심리-뇌영상 데이터 플랫폼을 중심으로 살펴보고자한다.


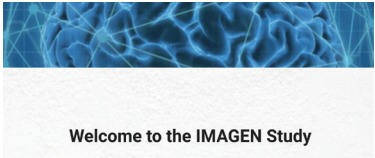

- 미국: 국립보건원 (NIH)의 기금으로 운영되는 ABCD (The Adolescent Brain Cognitive Development, ABCD study) 프로젝트는 미국 내 21개의 연구센터에서 9-10세의 아동 11,880명이 2013년부터 참여하였으며, 수집된 뇌영상 자료, 생체 표본, 행동 데이터는 미국 정신건강 연구소 데이터 아카이브(National Institute of Mental Health Data Archive, NDA)에 저장되어있다. 개방 과학 정신에 입각하여 2019년부터 해마다 업데이트를 진행하고, 일련의 허가 절차를 거쳐 연구자들에게 공유하고 있다.

- 캐나다: 캐나다 건강연구원 (Canadian Institutes of Health Research)과 캐나다 혁신재단 (Canadian Foundation for Innovation)의 기금으로 실행된 사그네 유년 연구(Saguenay Youth Study, SYS)는 2003-2015년에 시행되었으며, 퀘벡의 사그네 (Saguenay Lac Saint-Jean) 지역의 프랑스어를 모국어로 하는 캐나다인을 상대로 1,029명과 청소년과 그들의 부모 962명을 모집하여 플랫폼을 구축하였다. 세대 간의 유전자적 변이를 뇌, 체지방, 신장, 인지기능, 심혈관에 대한 데이터로 구분하여 구축 및 공유하고 있다.

- EU: 영국, 아일랜드, 프랑스, 독일의 연구소들이 동참하는 IMAGEN 프로젝트에서는 2,000명의 아동과 부모를 2-3년 간격으로 추적 연구하였다. 소아·청소년의 뇌발달과 행동, 특히, 보상 민감성, 충동성, 정서처리가 주요 연구 대상이며 수집된 데이터는 연구계획서를 제출한 연구자들에게만 공유하고 있다.

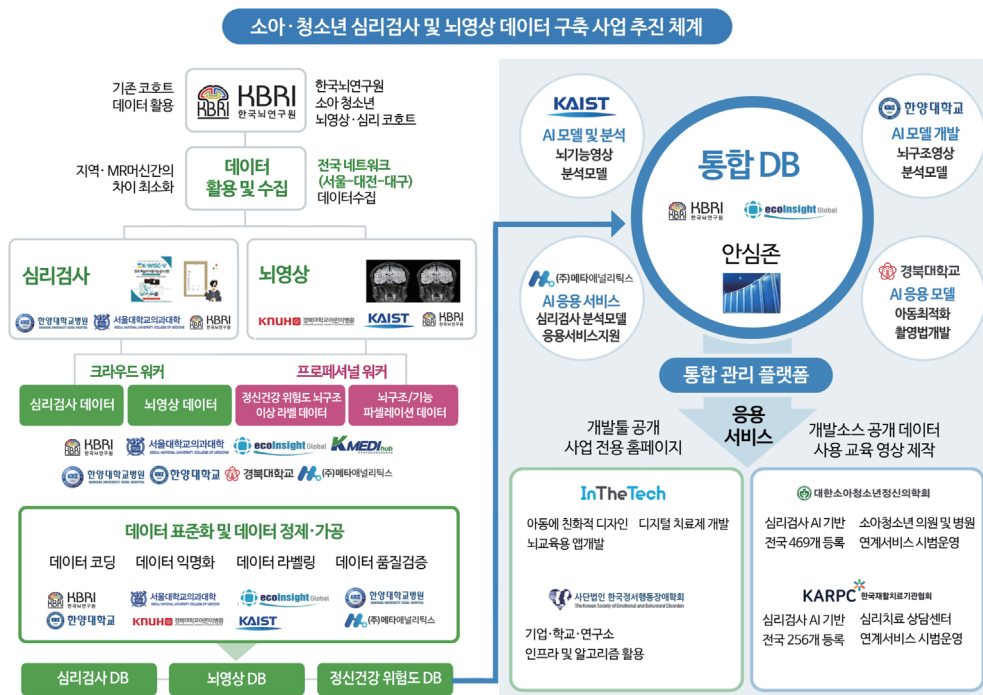
- 중국: 2013-2017년에 중국 쓰촨과 베이징에서 시행된 중국 색 둥지 프로젝트 (Chinese Color Nest Project, CCNP)는 6-20세까지의 1,520명의 건강한 개인을 피험자로 하고 있으며, 그 중 소아·청소년 수는 480명이다. 연구 데이터는 중국 과학 데이터 은행을 통해 공유될 예정이다.

- 일본: 2012년 시작된 도쿄 10대 코호트 연구(Tokyo TEEN Cohort study, TTC)의 확장연구인 뇌과학 도쿄 10대 집단 연구(population-neuroscience study of Tokyo TEEN Cohort, pn-TTC)는 심리, 인지, 사회, 신체적 측정값을 수집하고 뇌영상 기술과 접목하여 발달의 신경생리 기관을 규명하는 것을 목적으로 2년 간격으로 추적연구를 진행 중에 있다.

국가	특징	사이트
미국	<ul style="list-style-type: none"> <li>국립보건원과 뉴욕소재의 어린이 마음 연구소를 통한 소아·청소년 뇌영상 빅데이터 구축</li> <li>연구목적에 위해서만 공유 가능</li> </ul>	<p><a href="https://abcdstudy.org/">https://abcdstudy.org/</a></p> 
캐나다	<ul style="list-style-type: none"> <li>캐나다 건강연구원과 캐나다 혁신재단을 통한 소아·청소년 빅데이터 구축</li> <li>수집된 데이터에 뇌영상 데이터는 포함되어 있지 않으며 연구목적에 위해서만 공유 가능</li> </ul>	<p><a href="https://saguenay-youth-study.org/">https://saguenay-youth-study.org/</a></p> 
EU	<ul style="list-style-type: none"> <li>영국, 아일랜드, 프랑스, 독일의 연구소들이 동참하여 소아·청소년의 뇌발달과 정신건강 연구를 위한 빅데이터 구축</li> <li>연구목적에 위해서만 공유 가능</li> <li>플랫폼이 구축되어 있지 않음</li> </ul>	<p><a href="https://imagen-project.org/">https://imagen-project.org/</a></p> 
중국	<ul style="list-style-type: none"> <li>중국 색 동지 프로젝트를 통해 소아·청소년의 뇌발달과 정신건강 연구를 빅데이터 구축 예정</li> <li>수집된 데이터는 중국과학데이터 बैं크에서 공유 예정</li> </ul>	<p><a href="https://www.scidb.cn/en">https://www.scidb.cn/en</a></p> 
일본	<ul style="list-style-type: none"> <li>뇌과학 도쿄 10대 코호트 연구를 통해 소아·청소년의 뇌발달과 정신건강 연구를 위한 빅데이터 구축 예정</li> <li>수집 진행 중</li> </ul>	<p><a href="http://ttcp.umin.jp/">http://ttcp.umin.jp/</a></p> 

## 나. 소아-청소년 뇌연구 데이터 분석 플랫폼 현황

현재 한국뇌연구원을 중심으로, 서울대학교병원, 한국과학기술원(KAIST), 한양대학교, 한양대학교병원, 경북대학교, 경북대학교병원, 대구경북첨단의료산업진흥재단, 헬스케어 전문기업 에코인사이트, 메타에너지틱스와 함께 컴소시엄을 구축하였다. 이를 통해, 심리·뇌영상 데이터를 수집하고 심리적 정신건강도의 위험도를 라벨링 하여 심리검사 기반 정신건강 위험도를 AI로 예측하는 모델을 개발하고 있으며, 의료연계 플랫폼, 뇌영상 공유·통합 관리 플랫폼 운영 및 데이터 안심존을 구축할 예정이다.



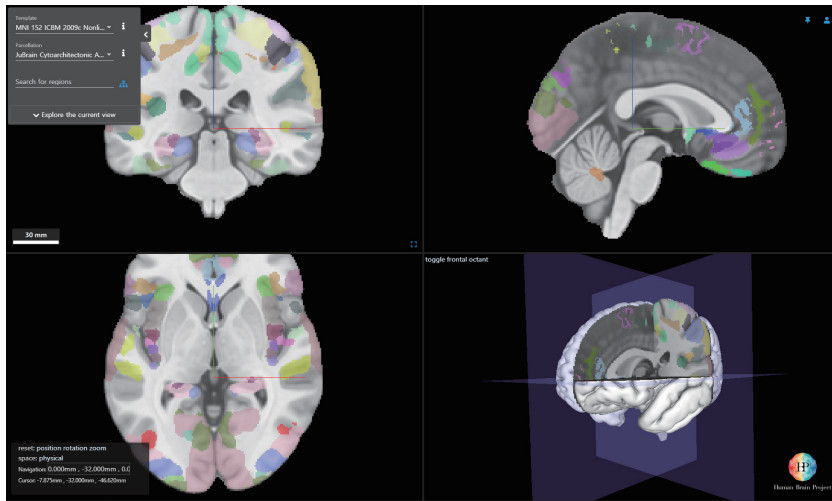
### 3

## 시각화 도구

### 가. BigBrain (<https://bigbrainproject.org/>)

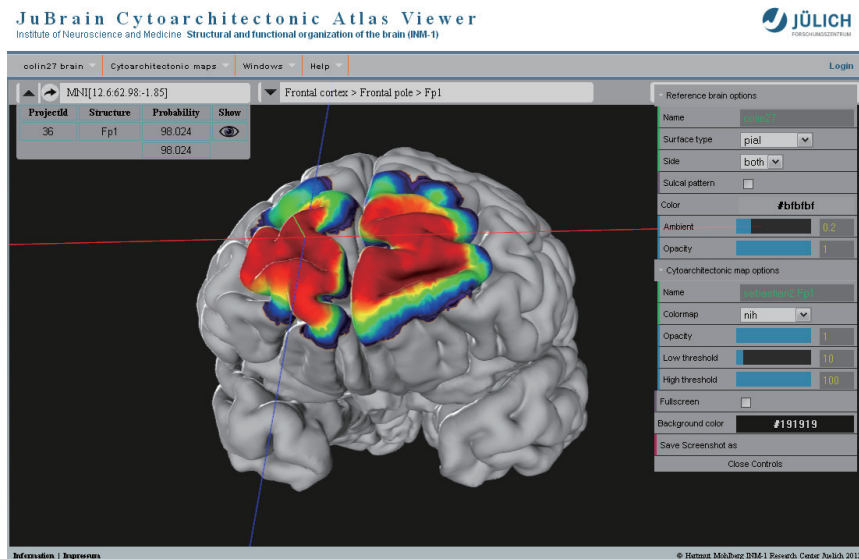
BigBrain은 2013년 6월 유럽 인간 두뇌 프로젝트의 일환으로, 캐나다 몬트리올 신경 연구소(Montreal Neurological Institute)와 독일의 율리히 연구소(Forschungszentrum Jülich)가 발표한 자유로운 열람이 가능한 고해상도 3D 디지털 인간 두뇌 지도 시각화 도구이다. 평균 1mm 해상도인 기존의 다른 인간 두뇌 3D 모델 시각화 도구보다 훨씬 미세한 20 $\mu$ m의 분해능을 지원하며, 3D Volumes, 3D Surfaces, 3D Classified Volumes 등의 정보를 제공하고 있다.





사용자는 뇌지도의 템플릿 공간과 참조구획을 선택할 수 있고, 뇌의 영역을 명칭으로 선택·검색하거나 화면에서 직접 지정할 수 있으며 뇌의 세부 영역별 다중 모드 데이터를 찾아 볼 수 있다. 대용량 데이터의 시각화를 위한 기술로 오픈 소스 프로젝트인 Neuroglancer를 사용하였다. 또한, 사용자는 원본 스캔 기능을 통해 시스템의 자체 분석 파이프라인을 사용한 결과값과 이로 부터 파생된 데이터에 이르기까지 다양한 데이터를 조회할 수 있다.

## 나. JuBrain Cytoarchitectonic Atlas Viewer (<https://julich-brain-atlas.de>)


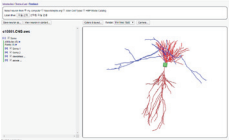
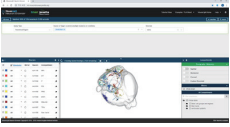
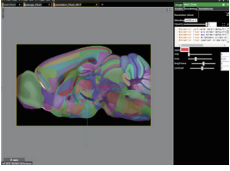


(<https://jubrain.fz-juelich.de/apps/cytoviewer/cytoviewer-main.php>)

JuBrain은 2002년 Juelich와 Duesseldorf의 Katrin Amunts와 Karl Zilles에 의해 개발되었고 Cytoarchitectonic 뇌지도를 시각화하는 시스템이다. Cytoarchitectonic 영역을 10개의 인간 사후 뇌의 조직학적 섹션에서 분석하였고, 이미지 분석 및 통계적 기준을 기반으로 면적 경계를 구역화하였다. 또한, 뇌과학 커뮤니티에 공개된 뇌지도를 사용할 수 있는 기능도 지원하고 있다. Occipital cortex, Parietal cortex, Temporal cortex, Frontal cortex, Insula 등 구역을 선택할 수 있는 다양한 옵션값을 제공하며 각 구역별 색상지정이나 구역 검색 등 사용자 맞춤형 시각화 옵션을 제공하고 있다.

## 다. 기타 시각화 도구

뇌연구 관련 빅데이터를 한눈에 볼 수 있게 도와주는 주요 시각화 시스템들을 아래에 요약하였다.

표 II-3	주요 뇌연구 빅데이터 시각화 시스템	
시스템	특징	화면
<p><b>Primate Cell Type Database</b></p> <p><a href="https://primatedatabase.com/morphology.html">https://primatedatabase.com/morphology.html</a></p>	<ul style="list-style-type: none"> <li>측면 전전두엽 피질로부터 샘플링된 인간 외 영장류(NHP) 세포의 세포내 기록을 시각화해줌</li> <li>전기 생리학, 형태학 및 분자(RNA 발현) 정보 등 뇌연구 실험실에서 생성된 일련의 실험 데이터를 제공</li> </ul>	
<p><b>HBP Neuron Morphology Viewer</b></p> <p><a href="https://neuroinformatics.nl/HBP/morphology-viewer">https://neuroinformatics.nl/HBP/morphology-viewer</a></p>	<ul style="list-style-type: none"> <li>SWC 및 Neurolucida(DAT, ASC, XML, NRX) 형식의 파일을 읽거나 NeuroMorpho, Allen, HBP Model Catalog 등의 데이터 베이스에 접속하여 데이터를 읽어들이는 다음 뉴런 형태로 시각화해줌</li> <li>사용자가 직접 뉴런을 보면서 편집할 수 있고 여러 형식의 파일로 저장하거나 확대하기 등의 다양한 기능을 제공</li> </ul>	
<p><b>MouseLight Neuron Browse</b></p> <p><a href="http://mouselight.janelia.org/">http://mouselight.janelia.org/</a></p>	<ul style="list-style-type: none"> <li>쥐의 뇌 연결성을 연구하는 신경 과학자를 위한 시각화 도구로 뇌 전체의 뉴런 재구성에 대해 심도있게 살펴볼 수 있음</li> <li>관심 있는 뇌 영역을 검색하고 수상돌기 또는 축삭 종점과 같은 특정 구조적 특징을 기준으로 검색하는 기능도 제공</li> </ul>	
<p><b>neurodata</b></p> <p><a href="https://neurodata.io/">https://neurodata.io/</a></p>	<ul style="list-style-type: none"> <li>동물과 인간의 지능 기전을 이해하는 것을 목표로 구축된 시스템이며, 뇌와 관련된 다양한 데이터와 분석 프로그램을 제공</li> <li>여러 연구 프로젝트로부터 얻은 MRI 데이터나 커넥톰(connectomes) 관련 데이터를 확인할 수 있고 그래프와 통계분석 도구를 이용할 수 있음</li> </ul>	



현재, 뇌연구의 세부 분야 간, 혹은 뇌연구와 타 연구 분야 간 융합연구의 중요성이 커지고 있지만, 이러한 연구방식을 지원할 수 있는 인프라는 아직 갖춰져 있지 않은 실정이다. 더불어, 뇌질환의 경우 조기진단에 따른 신속한 치료적 개입이 긍정적인 예후에 가장 중요한 요소이다. 이를 위해 대량의 뇌영상 데이터의 융합을 통한 인공지능 분석 알고리즘을 개발하여, 뇌질환 예방 및 조기 치료에 도움을 줄 수 있는 인프라가 필수적이다. 여러 세부 분야의 뇌연구 데이터들을 통합하여 빅데이터 기반의 분석과 시뮬레이션을 수행할 수 있는 인프라가 마련된다면, 뇌연구의 글로벌 경쟁력 확보와 높은 수준의 연구를 뒷받침할 수 있을 것으로 예상된다. 뇌연구 자원에 대한 체계적인 통합관리와 효과적인 융합연구를 위해서는 뇌연구자원에 대한 표준화된 데이터의 등록·관리부터 분석·시뮬레이션까지 수행가능한 통합 인프라 구축이 필요하다. 이 같은 통합 인프라를 통해 현재 각 소규모 연구 단위별로 분산되어 있는 뇌연구자원을 체계적으로 수집·등록·분석·활용하고 이중 데이터간의 통합분석과 연구자들간 협력할 수 있는 융합연구 환경을 마련할 수 있다.

## 참고문헌

1. 김보림, 융합연구정책센터, Weekly TIP - 뇌연구 국내외 정책동향 2018 January vol. 106.
2. 김영훈, 이준학, 정성진, 등, 공저, “뇌과학분야 빅데이터 분석을 통한 뇌질환 연구 혁신방안”, 한국과학기술정보연구원, 2015.
3. <https://www.geekwire.com/2022/allen-institute-to-create-atlas-of-whole-human-brain-with-100m-award-from-nih/>
4. Vanasse, Thomas J et al. “BrainMap VBM: An environment for structural meta-analysis.” Human brain mapping vol. 39,8 (2018): 3308-3325. doi:10.1002/hbm.24078.
5. Garrison JR, Done J, Simons JS. Interpretation of published meta-analytical studies affected by implementation errors in the GingerALE software. Neurosci Biobehav Rev. 2019 Jul;102:424-426. doi: 10.1016/j.neubiorev.2017.10.027. Epub 2017 Nov 10. PMID: 29103981.

## 참고문헌

6. Neuropixels data collection platform. Siegle, Joshua H et al. “Survey of spiking in the mouse visual system reveals functional hierarchy.” *Nature* vol. 592,7852 (2021): 86-92. doi:10.1038/s41586-020-03171-x
7. Vanasse T, Fox PM, Barron D, Robertson M, Eickhoff SB, Lancaster JL, Fox PT. BrainMap VBM: An Environment for Structural Meta-analysis. *Human Brain Mapping*, In-press, 2017
8. Mikula, S., Trotts, I., Stone, J., and Jones, E.G. 2007. Internet-Enabled High-Resolution Brain Mapping and Virtual Microscopy. *NeuroImage*. 35(1):9-15
9. Karcher, Nicole R, and Deanna M Barch. “The ABCD study: understanding the development of risk for mental and physical health outcomes.” *Neuropsychopharmacology : official publication of the American College of Neuropsychopharmacology* vol. 46,1 (2021): 131-142. doi:10.1038/s41386-020-0736-6.
10. Pausova, Zdenka et al. “Cohort Profile: The Saguenay Youth Study (SYS).” *International journal of epidemiology* vol. 46,2 (2017): e19. doi:10.1093/ije/dyw023.
11. Mascarell Maričić, Lea et al. “The IMAGEN study: a decade of imaging genetics in adolescents.” *Molecular psychiatry* vol. 25,11 (2020): 2648-2671. doi:10.1038/s41380-020-0822-5.
12. Gao, Peng et al. “A Chinese multi-modal neuroimaging data release for increasing diversity of human brain mapping.” *Scientific data* vol. 9,1 286. 9 Jun. 2022, doi:10.1038/s41597-022-01413-3
13. Okada, Naohiro et al. “Population-neuroscience study of the Tokyo TEEN Cohort (pn-TTC): Cohort longitudinal study to explore the neurobiological substrates of adolescent psychological and behavioral development.” *Psychiatry and clinical neurosciences* vol. 73,5 (2019): 231-242. doi:10.1111/pcn.12814.
14. Amunts, Katrin et al. “BigBrain: an ultrahigh-resolution 3D human brain model.” *Science (New York, N.Y.)* vol. 340,6139 (2013): 1472-5. doi:10.1126/science.1235381.
15. Henssen, Anton et al. “Cytoarchitecture and probability maps of the human medial orbitofrontal cortex.” *Cortex; a journal devoted to the study of the nervous system and behavior* vol. 75 (2016): 87-112. doi:10.1016/j.cortex.2015.11.006.



## 제3장 뇌연구자원 정보 연구 동향

**손태권** 한국뇌연구원 한국뇌은행 / taekwon@kbri.re.kr

손태권 박사는 부산대 생물정보학 박사(2006)를 마치고, 국가생명연구자원 정보센터 연구원(2005-2007), BML의원 연구원(2012-2015), 서울대 종합약학연구소 책임연구원(2019-2020)을 거쳐 현재 한국뇌연구원 한국뇌은행 선임연구원(2020-)으로 재직중이다. 연구분야는 생물정보학, 메타분석, 시스템 생물학, 빅데이터통합, 정보의학 등이다.

**김기범** 한국뇌연구원 연구전략실 / kpkim@kbri.re.kr

김기범 박사는 부산대학교 물리학 박사(2001), UCSB 박사 후 연구원(2006-2009)을 마치고, KAIST 물리학과 연구교수(2009-2014), IBS 첨단연성물질연구단 리서치 펠로우(2014-2016)를 거쳐, 현재 한국뇌연구원 연구전략실 선임연구원(2016-)으로 재직중이다. 연구분야는 단분자 생물물리, 생체신호 분석/활용 첨단연구장비 개발 등이다.

**장원준** 계명대학교 / mrdoin76@kmu.ac.kr

장원준 박사는 계명대 약학대학 박사(2015)를 마치고, 계명대 약학대학 박사 후 연구원(2015-2016), 약학연구소 연구조교(2016-2017)를 거쳐, 현재 계명대 약학대학 약학 연구소에서 연구조교(2017-)로 재직중이다. 연구분야는 중독, 약물학, 생물정보학, 시스템 생물학 등이다.

**이종민** 한양대학교 / ljm@hanyang.ac.kr

이종민 교수는 서울대 의용생체공학과 박사(1999)를 마치고, 한양대 의과대학 의공학 교실 조교수(2001-2004)를 거쳐, 현재 한양대 공과대학 전기제어생체공학부/바이오 메디컬공학전공 교수(2005-)로 재직중이다. 연구분야는 전신신경영상분석(의학영상 처리)이며, 특히 대뇌피질 모델링/분석, 뇌구조/기능영상 통합처리/분석, 뇌의 연결성 분석, 영상유전학 등이다.

### 목차 및 구성

#### 요약

1. 뇌영상/이미지 정보의 인공지능 분석
2. 오믹스 정보를 활용한 뇌질환 대상 신약 재창출
3. 뇌영상/이미지 정보와 오믹스 정보의 통합 분석
4. 주요 이슈 및 시사점

## 요약

지능정보를 근간으로 한 4차 산업혁명시대에 진입하면서 뇌연구자원 데이터 축적 기술과 다양한 뇌신경정보 분석 플랫폼 및 인공지능 기반 기술의 비약적 발전으로 뇌연구자원을 활용한 데이터 분석 기술 또한 새로운 도약의 전기를 맞이하고 있다. 뇌연구자원 정보 연구에는 뇌영상/이미지와 신경심리검사 데이터 분석에 초점을 맞춘 뇌신경정보학(neuroinformatics) 분야와 뇌신경과 관련된 오믹스 정보를 분석하는 생물정보학(bioinformatics) 분야 연구가 혼재되어 있으며, 최근에는 이 두 분야의 데이터를 통합 분석하는 연구가 활발히 이루어 지고 있다. 분석 대상 데이터의 규모도 점차 커지고 데이터의 종류가 다양해지고 있어서, 빅데이터 분석에 적용 되는 인공지능 기반 알고리즘의 적용도 활발히 이루어지고 있다.

## 1

## 뇌영상/이미지 정보의 인공지능 분석

뇌영상 데이터는 자기공명영상장치 (MRI) 기반의 T1 가중치 영상, 확산텐서영상(DTI, Diffusion Tensor Imaging), fMRI 등과 양성자단층촬영장치 (PET) 기반의 분자영상 등 다양한 종류를 포함하고 있다. 뇌영상데이터 분석은 전통적인 알고리즘 기반의 영상처리 및 기계학습에 기반하여 발전해왔으며 동시대 개발된 최첨단 방법론이 가장 먼저 적용되는 중요한 분야이기도 하다. 특히, 딥러닝을 중심으로 한 인공지능에 기반한 새로운 방법론의 대두는 뇌영상 데이터 분석에 있어서도 가장 중요한 방법론으로 자리잡고 있다. 뇌영상 데이터를 인공지능으로 분석하는 연구들은 뇌신경 발달성 장애, 퇴행성 장애 등의 다양한 뇌질환에 대한 조기 진단 예측 모델, 새로운 세부 분류 모델, 치료효과 검증 등을 포함하고 있으며 남녀간의 차이, 지능, 감정 등의 기전과 관련한 인지신경과학에서 다루는 중요한 주제들을 포함한다. 특히, 최근 활발히 진행되고 있는 인공지능 기반 치매 조기 진단 모델은 개인의 뇌영상을 입력 데이터로 받아 치매 발병 확률을 예측하게 하며, 예측된 치매 위험률은 치매 예방과 진단 등에 도움을 줄 수 있다. 뇌영상 빅데이터를 활용한 인공지능 치매 분류 모델은 전통적인 기계 학습 모델 대비 우수한 성능을 보여주고 있다. 기존의 전통적인 기계 학습 모델은 경도 인지 장애 환자 중 치매로 발전할 가능성이 높은 환자를 잘 분류하지 못하고 있으나, 심층 학습 모델을 사용하면 기존 알고리즘보다 더 높은 정확도로 경도인지 장애 환자 중 치매 위험군을 분류할 수 있다.

표 III-1

경도 인지 장애 분류 알고리즘 성능 비교

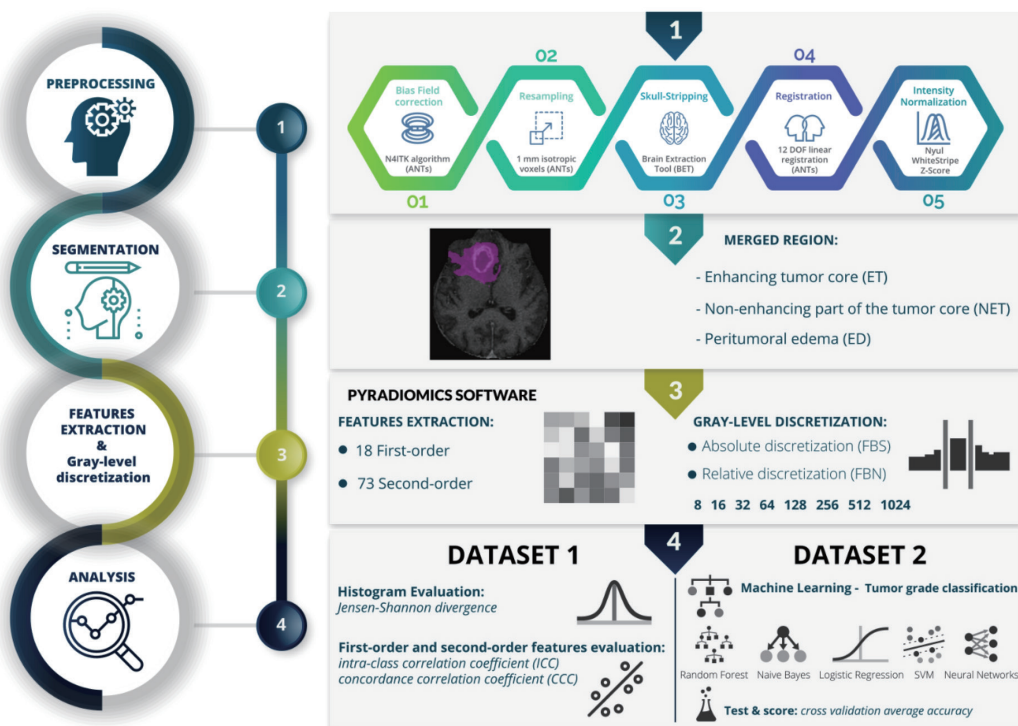
분류	방법론	데이터 종류	피실험 환자	ACC	SEN	SPE	AUC
Conventional methods (기존)	Gray <i>et al.</i>	$^{18}\text{F}$ -FDG PET	53pMCI+64sMCI	63.1	52.2	73.2	-
	Zhu <i>et al.</i> 2014	MRI, $^{18}\text{F}$ -FDG PET, CSF	43pMCI+56sMCI	70.9	42.7	94.1	77.4
	Zhu <i>et al.</i> 2016	MRI, $^{18}\text{F}$ -FDG PET	43pMCI+56sMCI	69.9	-	-	-
	Cheng <i>et al.</i>	MRI, $^{18}\text{F}$ -FDG PET, CSF	43pMCI+56sMCI	71.6	76.4	67.9	74.1
	Pan <i>et al.</i> 2019a	$^{18}\text{F}$ -FDG PET	166pMCI+360sMCI	79.43	69.14	84.16	83.88
	Pan <i>et al.</i> 2019b	$^{18}\text{F}$ -FDG PET	166pMCI+360sMCI	80.48	65.04	87.95	85.67
Emerging methods (신규)	Lu <i>et al.</i>	$^{18}\text{F}$ -FDG PET	112pMCI+409sMCI	82.51	81.36	82.85	-
	Suk <i>et al.</i>	MRI, $^{18}\text{F}$ -FDG PET	76pMCI+128sMCI	70.75	25.45	96.55	72.15
	Yee <i>et al.</i>	$^{18}\text{F}$ -FDG PET	210pMCI+427sMCI	74.7	74.0	75.0	81.1
	MiSePyNet	$^{18}\text{F}$ -FDG PET	166pMCI+360sMCI	83.05	72.12	88.06	86.80

표III-1은 기존의 경도 인지 장애 분류 알고리즘과 출처에 밝힌 논문 저자들이 개발한 심층 학습 모델의 성능을 비교하고 있다. 심층학습 모델이 네 가지 척도에서 더 우수한 성능을 보이는 것을 확인할 수 있다. 인공지능 연구의 성능을 결정짓는 가장 중요한 요소는 사용되는 데이터의 양과 질이라고 할 수 있기 때문에, 국내 연구의 경우 한국인의 데이터를 보다 적극적으로 활용하여 인공지능 연구를 수행하는 것이 중요한 점으로 작용하게 된다. 국내의 최근 연구에서는 한국인의 뇌에 최적화된 방법으로 학습하여 치매 분류 심층 학습 모델을 구축하였고, 사용자의 필요에 따라 미리 학습된 모델을 사용자가 가진 데이터를 이용해서 추가로 학습할 수 있도록 설계된 시스템이 개발되었다.

뇌영상을 이용하여 ICV (Intra Cranial Volume), GM (Gray Matter) 용적, WM (White Matter) 용적, 세분화된 뇌구역의 용적 등 다양한 뇌용적(volume)의 변화를 측정하여 대뇌의 발달 및 노화정도를 측정하는 방법 또한 활발히 연구되고 있다. 밝혀진 연구 결과들에 의하면 나이가 들어서 뇌 용적이 감소하는 것은 어쩔수 없지만, 뇌 구역별로 감소의 차이가 나타날 뿐만 아니라 개인별로도 차이가 존재한다는 것을 보여주고 있다. 뇌용적 뿐만 아니라, 대뇌피질의 두께, 뇌영역간의 연결성에 기반한 그래프를 이용한 분석들도 활발하게 연구가 이루어지고 있다. 특히, 인공지능 연구가 활발히 이루어짐에 따라 뇌용적 등을 계산하는 별도의 전처리 과정 없이 뇌영상 자체를 입력으로 받아 최종 결과를 제공하는 모델에 대한 연구도 중요한 연구 분야가 되고 있다. 이러한 뇌영상을 이용한 연구에 있어서도 데이터를 획득하는 장비간의 차이, 장비에 사용되는 파라미터의 차이, 개인간의 차이가 최종 결과에 큰 영향을 미치게 되며, 이를 극복하기 위해서는 이러한 변수들을 통제할 수 있는 적절한 수학적인 모형을 개발하거나 혹은 차이를 극복할 수 있는 대규모의 데이터를 활용하는 모델을 개발하는 것이 중요하다. 특히, 인공지능 기반 연구를 수행하기 위해서는 대규모의 데이터가 필수적인 요소가 되는데, 영상획득에 있어

서의 표준화, 데이터 포맷의 표준화, 데이터의 전처리 과정들의 표준화 등이 중요한 고려사항이다 [그림3-1]. MRI 기반의 다양한 유형의 영상들은 동일한 영상기법에 근거한 것이라고 하더라도 제조사에 따른 차이와 사용한 파라미터의 차이에 의한 획득영상간의 구조적인 차이가 무시할 수 없는 수준이기 때문에 뇌연구 자원 확보의 관점에서 획득 단계에서의 표준화는 매우 중요한 고려사항이다. 또한, 모든 주요 제조사는 DICOM 포맷 (<https://www.dicomstandard.org>)을 표준으로 획득된 영상을 제공하고 있으나, DICOM 포맷은 상대적으로 복잡하고 처리/분석 등의 목적으로 다루기에는 불편한 점이 있어 실제 연구에 있어서는 주로 NifTI 포맷 (<https://nifti.nimh.nih.gov>)을 사용하고 있다.

그림 3-1 인공지능 분석을 위한 뇌 MRI 영상의 표준화 과정의 예



출처 : Alexandre Carré, et al. (2020) "Standardization of brain MR images across machines and protocols", Scientific Reports 10: 12340.



## 2

## 오믹스 정보를 활용한 뇌질환 대상 신약 재창출

약물 표적 (drug target)은 일반적으로 약물이 결합하는 효소, 이온 통로, 수용체와 같이 약물과 결합할 수 있는 단백질을 의미한다. 약물은 체내에서 표적과 결합하여 유전자 조절에 영향을 주어 질병 치료에 작용할 수 있다. 따라서 질병을 치료하는 경로를 파악하기 위해서는 약물의 표적 정보를 수집하는 것이 매우 중요하다. 이러한 정보는 DrugBank, PubChem, KEGG DRUG, PharmGKB 등 공개약물 데이터베이스로부터 약물 표적 및 화학식, 분자 구조 등 다양한 생물학적 정보를 수집한다. 또한, 약물 부작용 정보는 SIDER와 FAERS (FDA Adverse Event Reporting System) 등의 데이터베이스를 많이 활용하고 있다.

그림 3-2

DrugBank 데이터베이스 (<https://go.drugbank.com>)

Drug Name	Interacting Gene/Enzyme	SNP(s)	Adverse Reaction	Reference(s)
5-Fluorouracil DB00544 (DRUGCARD)	Orotate phosphoribosyltransferase, Uridine 5-morphosphate synthase (Gene symbol = OPRT or UMPS) Swissprot P11172	<a href="#">rs1801019</a> (C Allele)	Neutropenia, diarrhea	Ichikawa W, Takahashi T, Suto K, Sasaki Y, Hirayama R: Orotate phosphoribosyltransferase gene polymorphism predicts toxicity in patients treated with bolus 5-fluorouracil regimen. Clin Cancer Res. 2006 Jul 1;12(13):3528-34. [PubMed]
5-Fluorouracil DB00544 (DRUGCARD)	Thymidylate synthase (Gene symbol = TYMS) Swissprot P04818	T8ER*2 <a href="#">rs34743033</a>	Neutropenia, diarrhea	Ichikawa W, Takahashi T, Suto K, Sasaki Y, Hirayama R: Orotate phosphoribosyltransferase gene polymorphism predicts toxicity in patients treated with bolus 5-fluorouracil regimen. Clin Cancer Res. 2006 Jul 1;12(13):3528-34. [PubMed]
5-Fluorouracil DB00544 (DRUGCARD)	Dihydropyrimidine dehydrogenase (Gene symbol = DPYD) Swissprot Q12852	<a href="#">rs1801265</a> (C allele) <a href="#">rs1801159</a> (G allele)	Nausea, vomiting, reduced white cell count	Zhang H, Li YM, Zhang H, Jin X: DPYD*5 gene mutation contributes to the reduced DPYD enzyme activity and chemotherapeutic toxicity of 5-FU: results from genotyping study on 75 gastric carcinoma and colon carcinoma patients. Med Oncol. 2007;24(2):251-8. [PubMed]
5-Fluorouracil DB00544	Glutathione S-Transferase P1 (Gene symbol = GSTP1)	<a href="#">rs1696</a> (A allele)	Hematological toxicity, gastrointestinal	Agostini M, Pasetto LM, Pucciarelli S, Terrazzino S, Ambrosi A, Bedini C, Galdi F, Friso ML, Mescoli C, Unso E, Leon A, Lise M, Nitti D: Glutathione S-Transferase P1 Ile105Val Polymorphism is Associated with Hematological Toxicity in Patients Treated

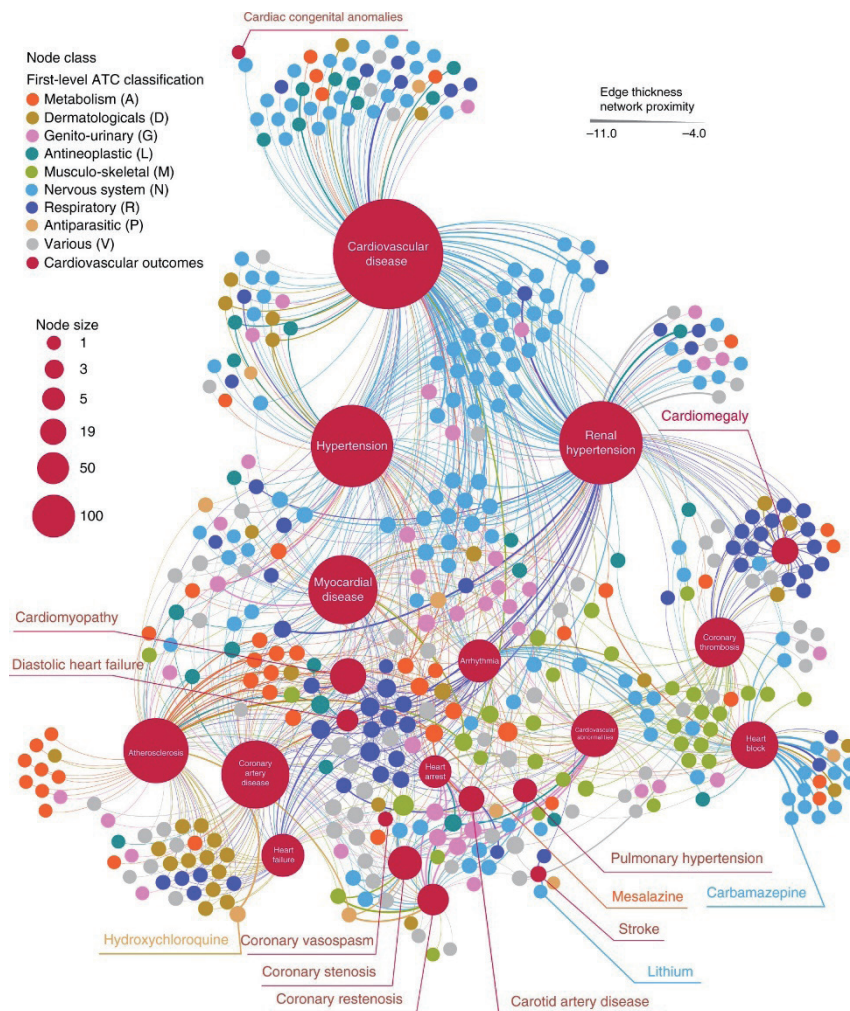
또한, 약물 반응에 의한 유전자 발현 데이터를 수집하기 위해 CMAP (Connectivity Map) 데이터베이스를 많이 활용하고 있다. CMAP은 약물 반응의 유전자 전사 데이터를 제공하는 데이터베이스로서, 이 데이터베이스는 1,300종 이상의 화합물을 다수의 세포에 적용한 7,000건 이상의 유전자 발현 프로파일을 보유하고 있다. 수집한 유전자 발현 프로파일의 차별 발현 분석 (differential expression analysis)을 통해 약물 반응에 의한 유의한 유전자군을 도출할 수 있다.



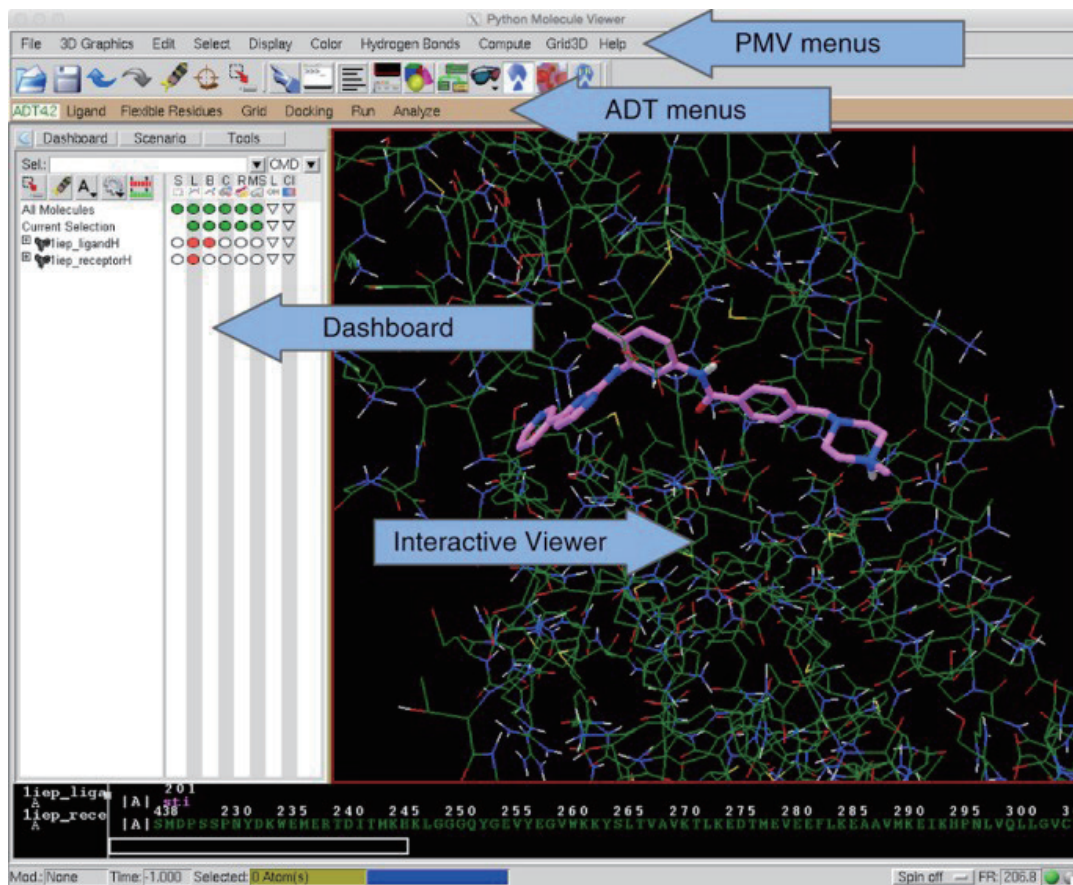
약물과 약물 표적의 생물학적 통합 네트워크 구축을 위해, KEGG, PID (Pathway Interactions Database), BioCarta, Reactome 등으로부터 약물표적의 대사경로 및 신호전달경로 정보를 수집하며 BioGrid, DIP (Database of Interaction Proteins), IntAct, HumanNet 등으로부터는 유전자 또는 단백질로 구성된 네트워크 정보를 수집한 후, 통합 네트워크로 구축할 수 있다.

다양한 데이터베이스로부터 수집한 데이터를 활용하여 약물재창출 후보 물질을 발굴하는 방법 중 대표적인 것으로 네트워크 기반의 약물간 기능적 유사도 분석 방법이 있다. 약물 간 유사도를 측정하여 약물의 새로운 적응증을 도출하는 방법은 신약 재창출을 수행하기 위하여 제안된 방법 중 가장 널리 사용되어 온 연구 방법이다. 이 접근법은 약물 간의 유사도가 높을 경우에는 서로 다른 두 약물이 비슷한 역할을 수행한다는 전제 조건을 기반으로 한다. 약물간 유사도 측정은 약물의 화학 구조식, 약물과 관련된 유전자, 염기 서열 등 약물과 관련된 다양한 정보를 기반으로 측정할 수 있다. 앞서 다양한 데이터베이스들을 통해 수집하여 구축한 통합 유전자 네트워크에서 표현된 약물 모듈을 활용하면 약물의 기능적 유사도를 측정할 수 있다. 약물 모듈의 유사도를 측정하기 위하여 군집 분석을 통해 모듈을 군집으로 적용하여 군집 간 거리를 기반으로 두 모듈 간의 유사도를 측정한다. 거리를 기반으로 측정하는 방법에는 네 가지 기준을 사용할 수 있다. 첫 번째는 유전자들 사이에서 발생할 수 있는 유전자 쌍의 거리 평균, 두 번째는 군집 유전자 쌍 중 가장 짧은 거리의 값, 세 번째는 유전자 쌍 중 가장 긴 거리의 값, 마지막으로 각 군집의 중심을 계산하고 그들간의 거리의 값을 측정하는 방법이다. 각 방법마다 장단점이 존재하기 때문에 다양한 거리 측정 방법을 적용 후, 최적의 군집 거리가 출현하는 군집 유사도 측정 방법을 선별하는 방식이 활용되고 있다.

다음으로, 네트워크 기반의 질병 간 기능적 유사도 분석을 통해 약물 재창출 후보 물질을 발굴하는 방법이 있다. 질병을 네트워크로 표현하는 접근 방법은 일반적인 질병 분류 체계에서 확인하기 어려운 질병 간의 연관 관계를 밝혀낼 수 있으며, 질병의 공통적인 원인과 특징을 식별할 수 있기 때문에 신약재창출을 수행하는 데 중요한 역할을 수행한다. 이러한 질병 마커를 이용하면 뇌질환과 다른 질병 간의 기능적인 유사도를 측정할 수 있다. 유사도 측정 과정에서 질병에 따른 유전자 발현 프로파일을 추가로 사용하면 유사도의 정확도를 향상시킬 수 있다. 먼저, 질병의 유전자 마커로 구성된 벡터를 형성한 후, 서로 다른 두 질병 간의 벡터 유사도를 측정하여 질병의 기능 유사도를 계산한다. 질병에 대한 유전자 발현 데이터는 질병 벡터를 구성하는 과정에서 실수 벡터로 표현한다. 즉, 유전자 발현량을 기반으로 질병 벡터를 구축하는 것이다. 두 질병 간의 상관관계를 측정하는 지표로는 피어슨 상관관계수 (Pearson correlation), 스피어만 상관관계수 (Spearman's correlation), 상호정보량 (Mutual information) 등을 활용하여 두 벡터 사이의 유사도를 측정할 수 있다.



위 과정에서 도출된 약물재창출 후보 물질은 생물정보학 분석을 통해 보다 가능성이 높은 후보 물질을 선별하는 과정을 거친다. 먼저, 생물정보학 데이터베이스 들을 활용하여, 약물재창출 후보 물질의 세포 내 작용위치, 단백질 약물 표적의 세포 내 위치 (subcellular localization), 상호작용 정보, 뇌혈관장벽(BBB, blood-brain barrier) 통과 여부 등을 확인하여 뇌질환 대상 약물재창출 가능성을 평가한다. 다음으로, 도킹 시뮬레이션 프로그램인 GOLD와 AutoDock Vina 등을 활용하여 단백질 표적과 약물재창출 후보물질의 결합 여부 및 결합력을 예측하여 최종 선별한다.



앞서 정량화 된 약물-질병-표적 관계 데이터를 컴퓨터가 이해할 수 있는 형태로 변환하고 기계 학습 과정에서 훈련 데이터와 검증 데이터로 분류한 후 훈련 데이터를 이용하여 모델을 구축하며, 검증 데이터로 모델의 유의성을 검증한다. 또한 기계 학습의 결과물로 기존에 알려지지 않은 약물-질병 관계를 도출 후 뇌질환 대상 신약 재창출을 위한 후보 약물군으로 식별한다.

인공지능을 활용하여 모델을 구축하기 위해, 전체 데이터 중에서 임의로 훈련 데이터와 검증 데이터로 나누게 되며, 훈련 데이터와 검증 데이터의 비율은 대부분의 연구에서도 사용 중인 10-fold를 기준으로 분류한다. 또한 분류 클래스는 알려진 약물-질병 관계를 제공하는 데이터 베이스인 CTD를 많이 활용하며 관계가 알려진 경우에는 참을, 그렇지 않은 경우에는 거짓을 부여한다. 분류 테이블의 각 특징은 본 연구를 통하여 계산된 약물과 질병의 유사도 값으로 구성 된다. 임의로 분류된 훈련 데이터를 입력 데이터로 사용하여 기계 학습을 통한 분류 모델을 구축한다. 분류에 사용되는 알고리즘에는 베이지안 분류 (Bayesian classification), 의사결정 나무 (Decision tree), SVM (Support Vector Machine), 신경망 알고리즘 등 다양한 방법이 적용되고 있다.



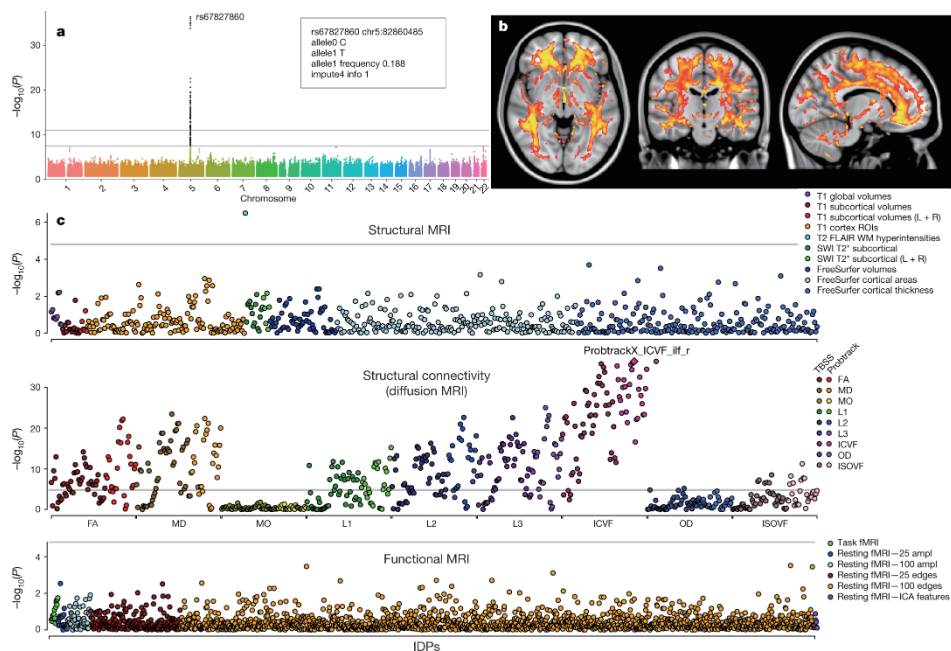
### 3

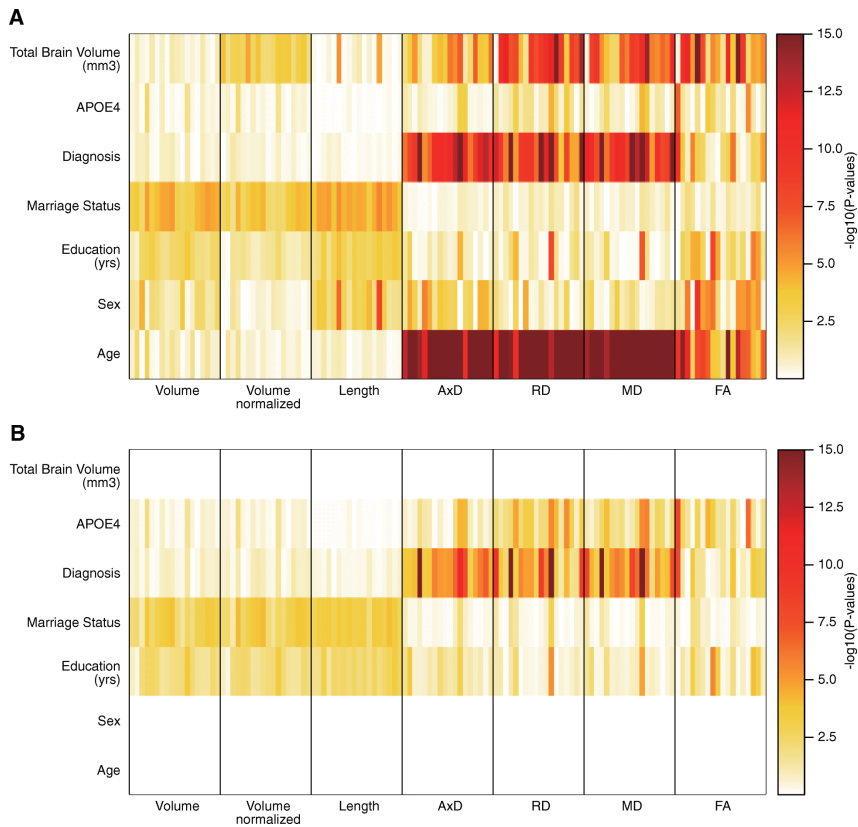
## 뇌영상/이미지 정보와 오믹스 정보의 통합 분석

최근 활발히 진행되고 있는 뇌영상/이미지 정보와 오믹스 정보의 통합 분석 연구에는 두 종류의 데이터간의 연관성을 분석하는 연구와 두 종류의 데이터를 함께 이용하여 질환 조기 진단 연구 등에 사용하는 두가지 유형이 있다. 방법론적인 관점에서는 통계적 방법과 기계 학습 방법이 함께 활용되고 있다. 주로, 회귀분석 (Multiple linear regression 등), 상관분석 (Pearson correlation 등), 매개 분석 등 기존의 통계적 기법으로 뇌데이터, 오믹스 데이터, phenotype 간 연관 관계를 도출하거나, 지도학습에 기반한 기계학습 기법 (Support vector machine, Random forest 등)을 적용하여 결과를 얻거나, 혹은 대규모 데이터를 이용한 딥러닝 등의 인공지능을 적용하는 사례가 점차 증가하고 있다. 한가지 사례는 Genomics 데이터 (SNP 등)과 건강보험공단 데이터 (시계열) 등을 활용하여 환자별 phenotype (처방 유형, 진단검사 수치, 예후) 중심으로 분석을 수행한 것이다. 또 다른 사례는 뇌영상을 endophenotype으로 활용하여 유전체 정보와의 연관성 분석을 수행한 것이다. UK Biobank 데이터에서 환자의 fMRI 데이터를 graph network 분석을 통해서 얻은 imaging markers를 endophenotype으로 설정하고 환자 혈액에서 genome/exome sequencing을 수행하여 SNP, CNV를 분석한 뒤, Multi-phenotype association test 등의 연관성 분석을 진행하였다 (그림3-5). 이 경우는, 질환명 등의 phenotype과의 연관성을 분석하는 것에 비하여 좀 더 직접적인 연관성 분석이 가능하다는 장점이 있다.

그림 3-5

유전체와 뇌영상의 연계분석 사례



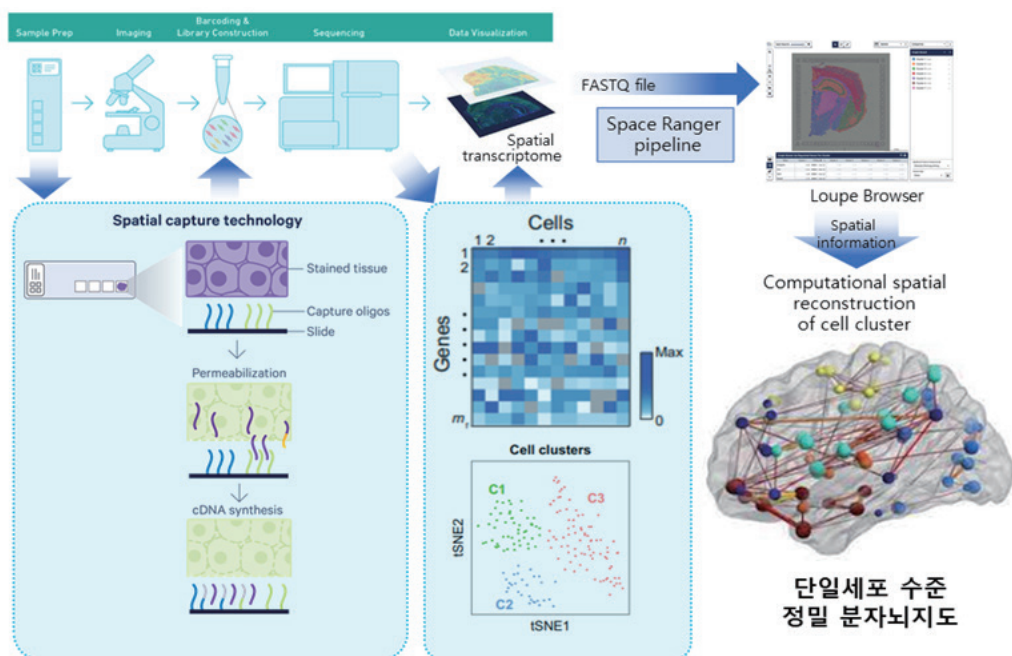


다음으로, 뇌정보와 전사체간의 통합 분석을 위해 MRI, phenotype, omics 간 상관관계 분석을 한 연구가 있다. 조울증 환자 및 정상인(healthy control)의 혈액의 전사체 및 MRI 데이터를 확보한 후, 환자의 MRI에서 gray matter와 white matter의 volume 변화량을 분석하고 구조(WBF, GMF, WMF, deep structures) 데이터에서 구조 비율의 변화량을 분석하였다. 이후 환자 혈액에서 전사체 서열 분석을 진행하여 차별발현 유전자(DEG, Differentially expressed gene)를 분석하였다. 통합 분석을 위해, 상관 분석(Pearson correlation) 및 Hamilton Depression Rating Scale, Young Mania Rating Scale등 과의 매개 분석(mediation analysis)을 진행하여, 뇌영상 데이터와 전사체, 표현형간의 상관관계를 분석한 결과가 보고되었다(그림3-6).

또한, 최근 뇌 특정영역의 전사체 발현 패턴을 단일 세포 수준으로 분석할 수 있는 공간전사체 분석 기술의 발달로, 분자수준의 뇌신경회로망을 분석을 통한 정밀 뇌지도 구축이 가능할 것으로 기대하고 있다(그림3-7).

그림 3-7

단일 세포 수준의 공간전사체 분석



뇌조직 이미지 데이터에서는 공간 정보를 확보하고, 단일 세포 수준의 좁은 영역에서 산출한 영역별 전사체 발현 패턴을 분석하여 뇌의 영역간 분자 발현 패턴의 상관관계와 연결성을 분석할 수 있다. 최근 배외측 전전두피질(dorsolateral prefrontal cortex)의 공간전사체 데이터가 보고되었고, 활발한 연구가 이루어지고 있다.

## 4

## 주요 이슈 및 시사점

최근 급격한 속도로 발전하고 있는 뇌영상/이미지 정보와 오믹스 정보를 활용한 분석 기술은, 뇌의 구조 및 기능상의 결함 등에 기인한 신체적, 정신적 질환 원인 규명과 이의 진단, 치료, 예방 기술에 가장 빠르게 적용되고 있다. 또한, 뇌신경계의 형성과 기능에 대한 생물학적 운영 원리를 규명하고, 이를 바탕으로 응용기술을 개발하는 뇌신경생물 분야에도 활발히 활용되고 있다. 그리고, 신경시스템이 외부로부터 정보를 받아들여 신경활동으로 전환하고 재구성하며 경험에 의해 변화하는 과정에서 형성되는 고등인지기능 연구와 그 응용기술을 개발하는 뇌인지 분야, 그리고 뇌의 구조와 기능 측정 및 모델링하는 기술을 바탕으로 뇌신경계와 외부기기를 융합하여 뇌신경계 기능을 회복하고 증진시키 기술을 개발하는 뇌공학 분야로 활용성을 넓혀가고 있다.

향후 뇌연구자원, 특히 뇌영상과 유전체정보를 이용한 뇌과학 연구는 인공지능을 적극적으로 활용하는 방향으로 전개될 것으로 예상된다. 이러한 전망하에서 다음 사항들을 특히 중요하게 다루어야 할 것으로 판단된다.

첫째, 대규모의 뇌영상 및 유전체 데이터 축적 및 이의 표준화이다. 의료영상은 데이터 획득 과정에서부터 환자 정보 보호 등의 이슈까지 훨씬 더 데이터 축적이 어려운 것이 현실이다. MRI의 경우 제조회사에 따른 영상결과물의 차이, 영상화에 사용하는 파라미터에 따른 영상결과물의 차이 등이 다기관이 참여하는 형태의 데이터 획득에 큰 문제점을 야기하게 된다. 뇌영상 및 유전체 데이터 확보에 필요한 경제적 사회적 비용을 고려하면 획득 단계에서부터 표준화를 위한 적극적인 노력이 필요하다. 세계적인 표준화 작업에 효과적이고 주도적인 대처를 위해 공신력 있는 기관 주도로 최대한 많은 연구집단의 참여를 이끌어내는 방식이 바람직하다.

둘째, 국가에서 위임받은 기관이 하나의 통일된 형태로 데이터를 관리하고 이를 국내외 연구자들에게 효율적으로 제공할 수 있는 공개플랫폼의 구축이다. 데이터를 분산하여 각각의 기관이 관리하면서 데이터에 키워드를 붙이고 이를 검색하여 활용할 수 있도록 하는 방안은 의료영상 특히 뇌영상 및 유전체 데이터에는 적합하지 않은 모형이며, 하나의 거대한 데이터베이스 시스템 하에서 여러 유형의 데이터들이 개인을 중심으로 유기적으로 잘 묶여져 있고 이를 손쉽게 활용할 수 있도록 하는 것이 무엇보다 중요한 점으로 생각된다. 또한, 개인 정보 뿐만 아니라 질환과 관련된 민감한 정보들이 포함된 데이터이기 때문에 일반인에 대한 데이터 개방과 관련해서도 적절한 논의 및 감시감독이 필요하다.

셋째, 뇌연구자원 데이터를 일반 연구자들에게 공개했을 때, 이러한 데이터를 이용한 결과물을 공개하는 것도 중요하지만 특히 학습된 인공지능 모델의 공개가 중요하다. 인공지능 연구에 있어서 방법론의 우수성도 중요하지만 본질적으로 활용되는 데이터가 훨씬 더 큰 역할을 수행하기 때문에, 분석의 재현성이나 개선방안 마련 등을 위해 데이터 분석에 이용한 인공지능 모델도 공개하는 것이 바람직하다고 생각된다. 물론, 사업화 등과 관련해 일정 기간 비공개 등록을 하는 등에 대한 논의도 중요한 부분이며 충분한 논의와 합의점 도출이 필요하다.

넷째, 인공지능 기반 연구에 있어서 더 많은 데이터 확보에 대한 노력 외에도 방법론에 대한 고찰도 중요할 것이다. 지도학습 방법 뿐만 아니라 사람이 정답을 붙여놓지 않은 상태로 현장(병원 혹은 연구소)에서 생성되고 있는 대부분의 기존 데이터를 이용하기 위한 비지도 학습 방법의 적극적 활용이 가능해야 할 것이다. 그리고 적은 수의 데이터를 이용한 메타 학습방법의 개발 등도 향후 많은 연구가 이루어져야 할 것이다.

세계적으로 빅데이터와 첨단 장비 및 기술을 활용한 혁신적인 융합기술 구현이 이루어지고 있으며 이에 따라 대규모 연구프로젝트들이 빠르게 진행중이다. 또한, 연구발전의 빠른 속도만큼 뇌 관련 산업시장이 성장하고 있다. 국내에서 뇌연구촉진법 개정 등 다양한 뇌연구 육성정책을 통해 인력양성 및 연구경쟁력을 확보하고 세계 수준의 연구성과가 도출되고 있으나, 추격형 연구 패러다임의 극복에는 한계가 있다. 선택과 집중을 통한 R&D 고도화와 뇌연구 생태계의 내실을 다져 뇌연구 선진국으로 도약하기 위한 노력과 경주가 절실한 시점이다.



---

---

## 참고문헌

1. Pan, Xiaoxi, et al. "Multi-view separable pyramid network for AD prediction at MCI stage by 18 F-FDG brain PET imaging." *IEEE Transactions on Medical Imaging* 40.1 (2020): 81-92.
  2. Elliott, Lloyd T., et al. "Genome-wide association studies of brain imaging phenotypes in UK Biobank." *Nature* 562.7726 (2018): 210-216.
  3. Woo, Young Jae, et al. "Comparison of brain connectomes by MRI and genomics and its implication in Alzheimer's disease." *BMC medicine* 18.1 (2020): 1-17.
  4. Anand, Amit, et al. "Integrative analysis of lithium treatment associated effects on brain structure and peripheral gene expression reveals novel molecular insights into mechanism of action." *Translational psychiatry* 10.1 (2020): 1-10.
  5. Eshraghi, Mehdi, et al. "Enhancing autophagy in Alzheimer's disease through drug repositioning." *Pharmacology & Therapeutics* (2022): 108171.
  6. Arbo, Bruno Dutra, et al. "Repositioning and development of new treatments for neurodegenerative diseases: Focus on neuroinflammation." *European Journal of Pharmacology* (2022): 174800.
  7. Agostini, Francesco, et al. "Metformin repurposing for parkinson disease therapy: Opportunities and challenges." *International Journal of Molecular Sciences* 23.1 (2021): 398.
  8. Wishart, David S., et al. "DrugBank: a comprehensive resource for in silico drug discovery and exploration." *Nucleic acids research* 34.suppl\_1 (2006): D668-D672.
  9. Eberhardt, Jerome, et al. "AutoDock Vina 1.2. 0: New docking methods, expanded force field, and python bindings." *Journal of Chemical Information and Modeling* 61.8 (2021): 3891-3898.
  10. Cheng, Feixiong, et al. "Network-based approach to prediction and population-based validation of in silico drug repurposing." *Nature communications* 9.1 (2018): 1-12.
  11. Forli, Stefano, et al. "Computational protein-ligand docking and virtual drug screening with the AutoDock suite." *Nature protocols* 11.5 (2016): 905-919.
  12. Carré, Alexandre, et al. "Standardization of brain MR images across machines and protocols: bridging the gap for MRI-based radiomics." *Scientific reports* 10.1 (2020): 1-15.
  13. 의료기기산업팀, 한국보건산업진흥원, "인공지능(AI) 기반 의료기기 현황 및 이슈", 보건산업브리프 - 의료기기 주요 이슈 분석, 2018.
  14. Maynard, Kristen R et al. "Transcriptome-scale spatial gene expression in the human dorsolateral prefrontal cortex." *Nature neuroscience* vol. 24,3 (2021): 425-436. doi:10.1038/s41593-020-00787-0
- 
-



---

## 편집위원

서판길	KBRI/원장	백승태	포항공대 생명과학과/교수
정윤하(위원장)	KBRI/뇌연구정책센터장	최형진	서울대학교 의과대/교수
이계주	KBRI/전략실장	김성필	UNIST 인간공학과/교수
민병철	KBRI/경영기획 실장	임창환	한양대학교 생체공학과/교수
이장재	한국과학기술기획평가원(KISTEP)/ 혁신전략연구소 소장	차인준	KBRI/연구원(편집간사)

『Brain Insight』의 내용 인용 시 반드시 <뇌연구 빅데이터 현황 및 연구 동향>이라는 출처를 명시해주시기 바랍니다.

『Brain Insight』의 내용에 대한 문의사항은 한국뇌연구원 뇌연구정책센터(T. 053.980.8516)로 연락주시기 바랍니다.



2022 Vol.05 (통권 제5호)

---

발행일 : 2022년 11월

발행처 : 한국뇌연구원 뇌연구정책센터  
(41062) 대구광역시 동구 첨단로 61  
T. 053.980.8516 F. 053.980.8519  
<http://www.kbri.re.kr>

인쇄처 : (주)드림디앤디(T. 02.2268.6940)

