

국가 연구데이터 개방 및 활용 활성화 방안

2020. 9. 4.

KISTI 이상환



1

배경 및 필요성

2

해외 동향

3

국내 동향

4

한계와 문제점

5

연구데이터 개방과
공유 활성화 방안

6

결론



1. 배경 및 필요성(1/2)

연구데이터는 국가 과학기술 경쟁력을 좌우하는 원천이며, 연구개발 방식도 **데이터 기반으로 전환**

데이터 가치 증대

데이터의 활용이
과학기술 경쟁력의 핵심으로 대두

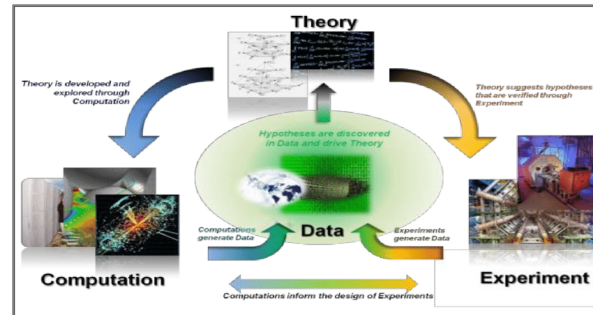
- 4차 산업 혁명은 인공지능과 함께 빅데이터가 주요한 변화 동인
- 데이터의 분석 및 활용으로 새로운 과학적 발견과 지식의 창출로 고부가가치화



연구개발 패러다임 변화

디지털 기술의 발전에 따라
데이터 기반의 R&D 전환

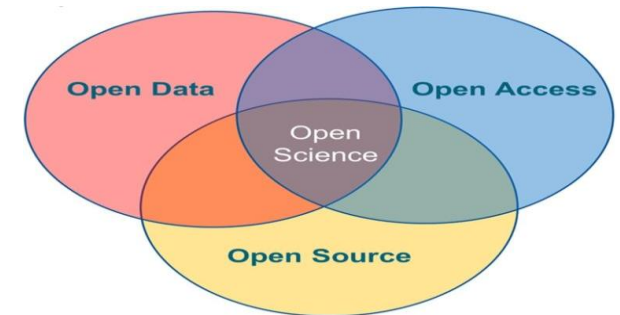
- 디지털 기술의 발전은 연구자의 연구 환경을 변화시키고 있음
- 연구의 방식도 데이터 중심으로 연구개발 패러다임이 변화하고 있음



오픈사이언스

공적 자금을 통해 생산된
연구결과에 대한 쉬운 접근과 재 활용

- 오픈 액세스 저널, 온라인 논문 공개 증가 등 개방, 공유 사례 증가
- 연구개발 과정 및 성과의 개발·공유를 국가적 차원에서 추진



1. 배경 및 필요성(2/2)

연구데이터 공유와 활용은 연구결과의 신뢰도 제고 등 연구개발 경쟁력과 효율성 강화

R&D 효율성 제고

연구 활용성 제고 및 중복 연구 최소화

- 연구데이터 수집, 가공, 처리 등에 투자되는 시간과 비용을 절감하여 연구개발 주기를 획기적으로 단축
- 필요한 연구데이터를 재활용함으로써 투자 효율성 증대



연구개발 경쟁력 향상

연구 재현성 확보 및 융합 연구 확산

- 연구데이터 공개를 통해 연구 재현성 보장 및 투명성 확보
- 데이터 기반 커뮤니티 형성으로 학제 간 융합 및 공동 연구 활성화

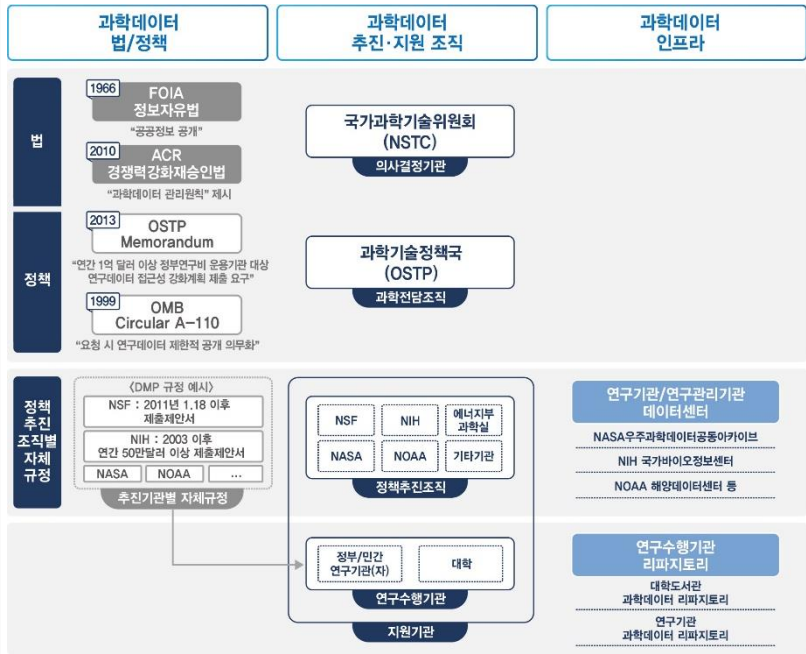


2. 해외 동향

주요 선진국은 국가 차원에서 연구데이터 공유 · 활용 정책을 시행하고 인프라 구축 · 운영하고 있음

미국

연방기관 중심으로 연구데이터 공유 정책 추진
(연간 1억 달러 지출하는 연방기관 대상)



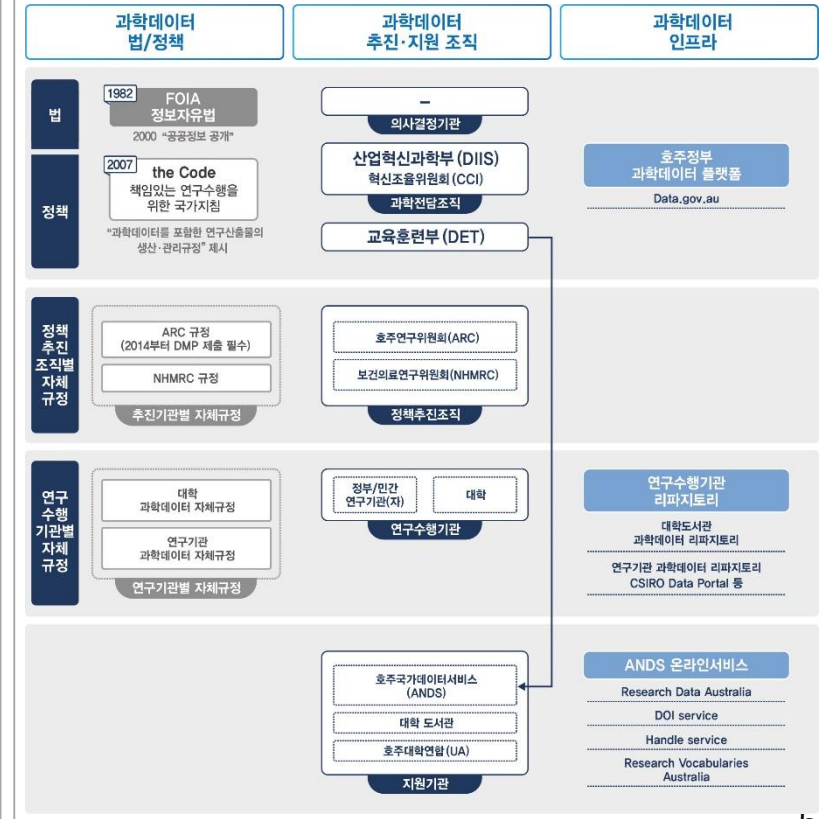
영국

연구회 중심으로 연구데이터 공유 정책 추진
(7개 연구 이사회)



호주

연구회 중심으로 연구데이터 공유 정책 추진
(영국과 유사한 추진 체계)



3. 국내 동향 - 정부 정책 수립

연구데이터에 대한 국가 차원의 관리와 활용에 대한 국가 정책 및 의지 표명

비전 연구데이터 지식 자산화 및 빅데이터화로 혁신 성장을 선도하여 국가연구개발 효율성 및 경쟁력 강화

목표	국가 연구데이터 공유 활성화			
	연구데이터 공유·활용의 인식제고	다양한 분야들을 결합한 융합 연구 활성화	연구 효율성 및 투명성과 연구결과물의 수준 제고	연구데이터 공유·활용을 위한 전문인력 양성

전략	1 융합 연구 및 오픈사이언스 등 선진 연구 환경 조성	2 계층적 공유·활용을 위한 추진체계 구축	3 데이터 공유·활용의 가치를 단기간에 제시할 시범사업 추진	4 융합공동연구 역량 강화를 위한 분야별 연구 공동체 형성	5 연구데이터 관련 분야별 전문인력 양성
-----------	--	-----------------------------------	---	--	----------------------------------

* 국가 연구데이터 공유·활용 전략안('18. 1, 국과심 안건 상정 및 통과)

3. 국내 동향 - 법적 근거 및 기반 마련(1/2)

연구데이터 정의와 데이터관리계획(DMP) 도입 및 DMP정보 등록 근거 신설('19. 9. 1. 시행)

국가연구개발사업의
관리 등에 관한 규정

입법예고
('18.12.26)

공고
('19. 3. 19)

시행
('19. 9. 1)

<p>1 정의 및 DMP 도입근거</p>	<p>연구데이터 정의(2조 18항 신설) DMP 정의(2조 19항 신설) 연구데이터 생산·보존·관리 및 공동활용 시책 수립·추진(25조 28항 신설)</p>
<p>2 DMP 제출</p>	<p>연구개발계획서 작성 시 DMP 제출 요구(6조 4항 개정) 연구개발과제의 선정 시 검토 요구(7조 3항 개정) 과제 협약 체결 시 DMP 제출 요구 및 준수(9조 1항 개정)</p>
<p>3 DMP 보고</p>	<p>최종보고서 작성 시 DMP 제출 추가 요구(15조 2항 개정)</p>
<p>4 DMP 관리 및 연계</p>	<p>DMP에 포함된 정보 등록 및 연계(25조 16항 개정)</p>

3. 국내 동향 - 법적 근거 및 기반 마련(2/2)

과기정통부 산하기관(NST, IITP, NST)의 DMP 적용을 위한 이행방안 마련('18.12) 및 추진('19.1)

연구재단

- 2019년
 - 바이오 분야 7개 사업 58개 과제
 - 나노소재 분야 1개 사업 25개 과제
- 2020년
 - 바이오 분야 6개 사업 82개 과제
 - 나노소재 분야 7개 사업 82개 과제
 - 다부처 분야 1개 사업 5개 과제
 - 반도체 분야 5개 사업 39개 과제

정보통신기획평가원

- 2019년
 - ICT 분야 11개 과제
- 2020년
 - ICT 분야 7개 과제

국가과학기술연구회

규정 개정 후 시범 적용

- 연구회의 “소관기관 주요사업 관리규정” 개정 후 DMP 시범적용

연구회(NST)

소관기관 주요사업 관리 규정

출연(연)

주요사업 관리 규정

* 13개 출연(연) DMP 적용('20.5.3)

3. 국내 동향 - 기반 인프라 구축

연구데이터 공유 및 활용을 위한 국가연구데이터플랫폼(DataOn) 개발·운영('20. 1)



4. 한계와 문제점

연구성과물 미지정으로 국가 자산 관리 한계

- 현행 R&D 법률상 연구데이터는 연구성과물로 지정되어 있지 않아 국가 차원에서 관리 한계
- 연구성과로서의 평가, 관리, 유통의 대상이 아님

연구데이터 소유와 공개에 대한 연구자 인식 한계

- 국가 R&D 성과는 연구자 자신의 것으로 인식
 - 연구데이터도 기존 성과물처럼 소유와 저작권을 분리
- * (현행) 연구개발 성과물에 대한 관리 주체는 기관이고, 연구자는 저작권 등 권리 보장

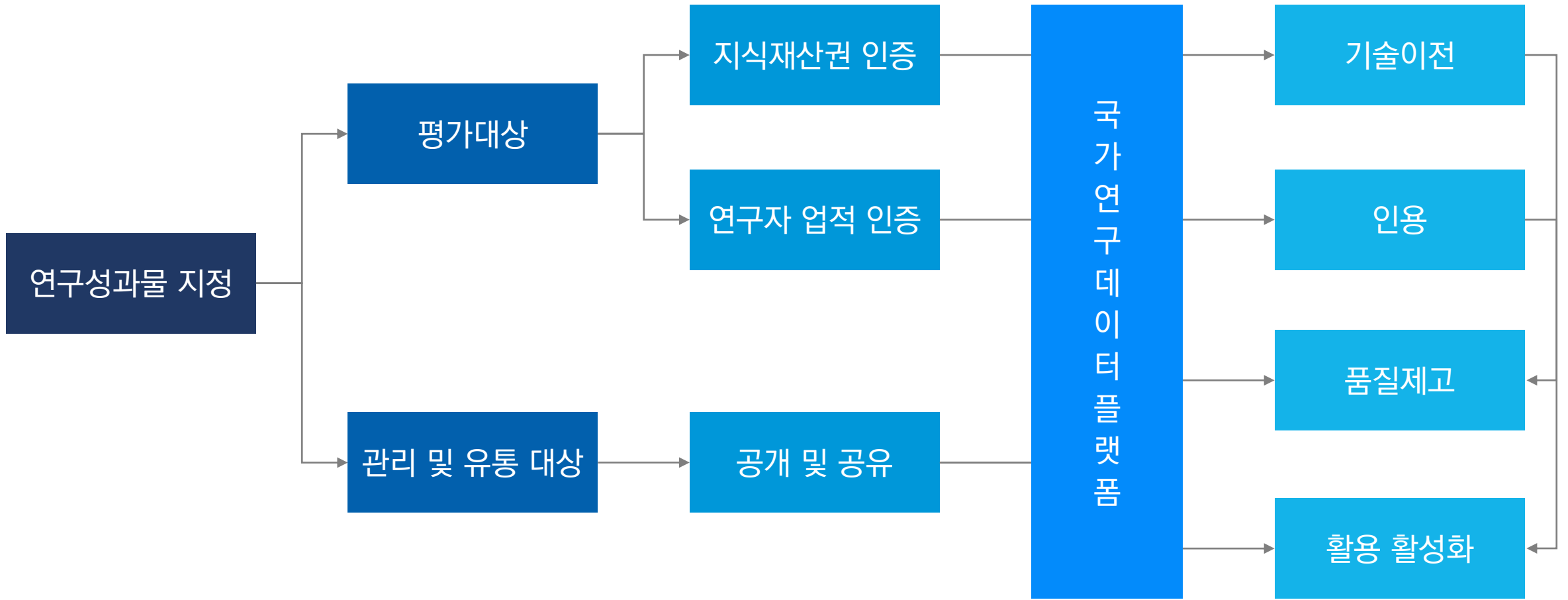
연구자가 자발적으로 참여하는 체제 미흡

- 연구자가 연구데이터를 제출하고 공개하는 데 안심하고 혼란이 없도록 가이드라인 부재
- 연구자가 자발적인 참여를 통해 연구데이터의 품질을 높이고 선순환적인 생태계를 높이는 인센티브 부재

연구데이터 관련 제도적 기반 후퇴

- (국가) 국가연구개발혁신법과 시행령에 연구데이터 관련 조항 부재로 연구데이터 관리와 활용 근거 미약
- (기관) 연구자 개인이 연구데이터 관리 주도하여, 기관 차원에서 연구데이터 관리와 활용에 한계 발생
- (연구자) 많은 연구자들이 다른 연구자들의 연구데이터를 필요로 하나, 연구데이터 존재와 보유기관 확인 어려움

5. 연구데이터 개방과 활용 활성화 방안 – 전체적인 개념



5. 연구데이터 개방과 활용 활성화 방안 – 정책 및 제도(1/3)

연구성과물 지정

- 국가R&D 사업 관련 법률에 연구데이터 정의와 더불어 연구성과물로 지정
 - 국가의 자산으로 관리, 유통, 평가 대상으로 근거 마련

전담기관 및 시스템 지정

- 연구데이터의 관리 및 활용을 위해 전담기관과 전담시스템 지정
 - 과제정보-10대 성과물-연구데이터 상호 연계를 통해 국가R&D 전주기별로 연구자에게 과학기술지식정보와 데이터를 제공

DMP 제도 의무화

- 국가연구개발혁신법 시행령에 DMP 제도를 의무사항으로 반영
 - 국가 차원에서 연구데이터 관리와 공유에 대한 근거 제시
 - 현행 공동관리규정은 중앙행정 기관장이 필요시 적용

* DMP(Data Management Plan)
R&D를 통해 생산되는 데이터를 기술하는 계획서

5. 연구데이터 개방과 활용 활성화 방안 – 정책 및 제도(2/3)

자발적인 공개 및 활용에 대한 기본원칙 제시

- 규제와 책임과 관련된 내용은 최소화하되, 연구자의 연구성과 공개 의무를 기본원칙으로 제시
 - 과제별 특성에 따라 데이터 공개방법 및 시기 등에 유연성 부여하나, 표준적인 가이드라인 제시
- 연구데이터 공개 및 활용에 따른 연구자 권익보호
 - 공개 시점과 공유방식을 연구자가 스스로 정하게 함으로써 연구자의 권리를 보호
 - 자발적인 공개 및 활용 활성화를 위한 인센티브 방안 마련

연구 '자율성'에 따른 연구자의 '연구윤리' 강조

- R&D 결과물인 연구데이터의 공개를 권고하고, 기본적인 연구자의 연구윤리와 의무사항으로 강조
 - 연구데이터는 과학연구의 객관성과 진실성을 검증하기 위한 중요한 자료이며, 연구데이터에 대한 품질과 신뢰성에 대한 1차적인 책임은 연구자임을 강조
- 특히, 연구데이터 품질과 신뢰성은 해당 분야의 전문가 아니면 품질 검증의 한계가 있어, 공개를 통한 품질 및 신뢰성 강화

5. 연구데이터 개방과 활용 활성화 방안 – 정책 및 제도(3/3)

연구자의 면책조항 마련

- 연구데이터 활용으로 인해 문제가 발생한 경우
 - 연구데이터 생산한 연구자의 면책조항, 이의신청 근거 및 절차 명확화
- 연구자의 연구결과 실패 인정
 - 실패의 원인과 연구데이터를 함께 공개하여 반복적인 실패 연구를 미연에 방지

연구자의 연구 업적 인정

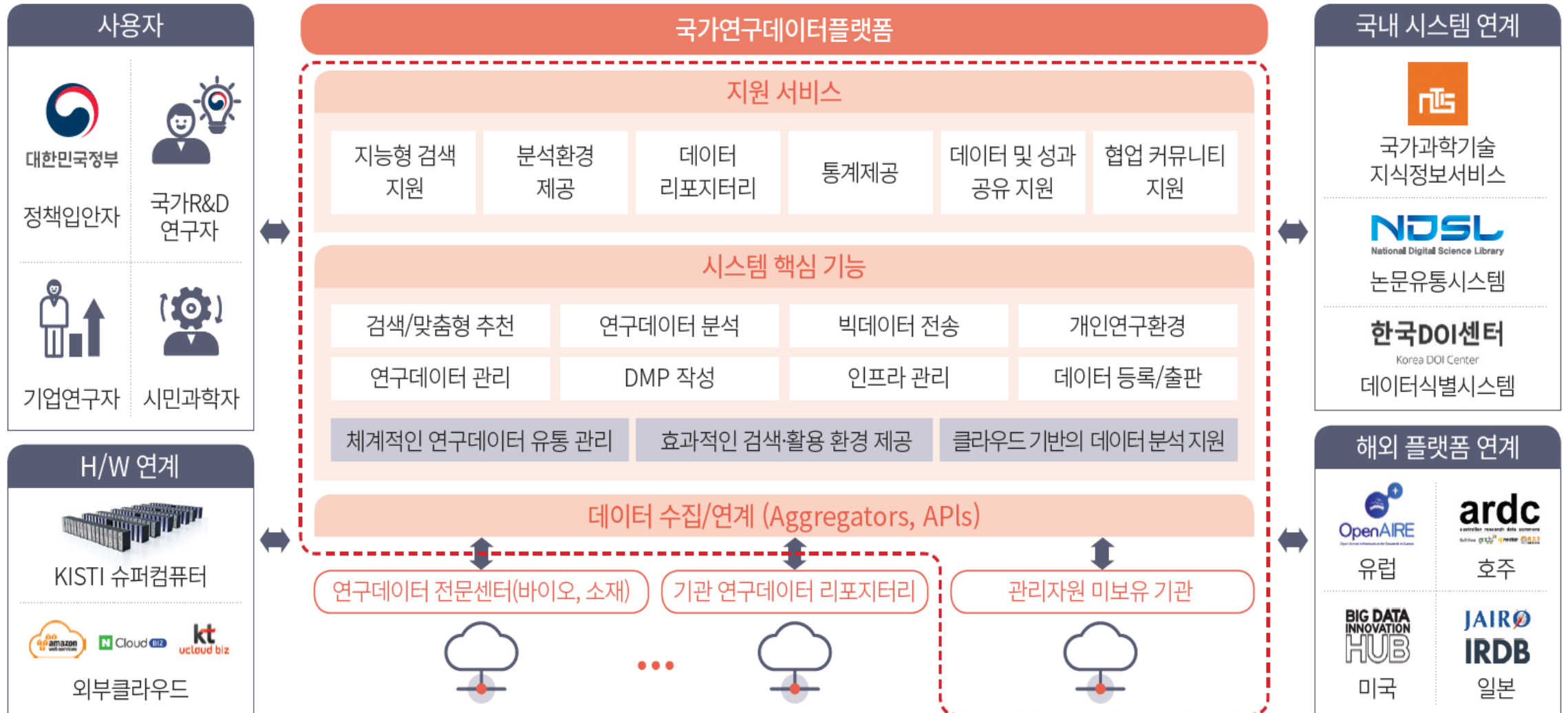
- 연구데이터 인용 기반 마련
 - 연구자가 생산된 연구데이터를 등록할 경우, DOI(Digital Object Identifier) 번호를 부여
- 표준적인 연구데이터 인용 방식 마련
 - 논문이나 연구보고서에서 데이터를 인용할 경우, 논문과 같은 성과로 인정하고 데이터 활용에 대한 추적 관리

연구데이터 보상 제도 마련

- 연구 목적으로 연구데이터 활용
 - CCL(Creative Commons License) 등과 같은 공개 라이선스 활용
- 상업 목적으로 연구데이터 활용
 - 특허와 같은 기술이전 방식으로 연구자에게 금전적인 보상체계를 마련하여 제공
 - * 기관 50%, 연구자 50% 배분

5. 연구데이터 개방과 활용 활성화 방안 – 인프라(1/2)

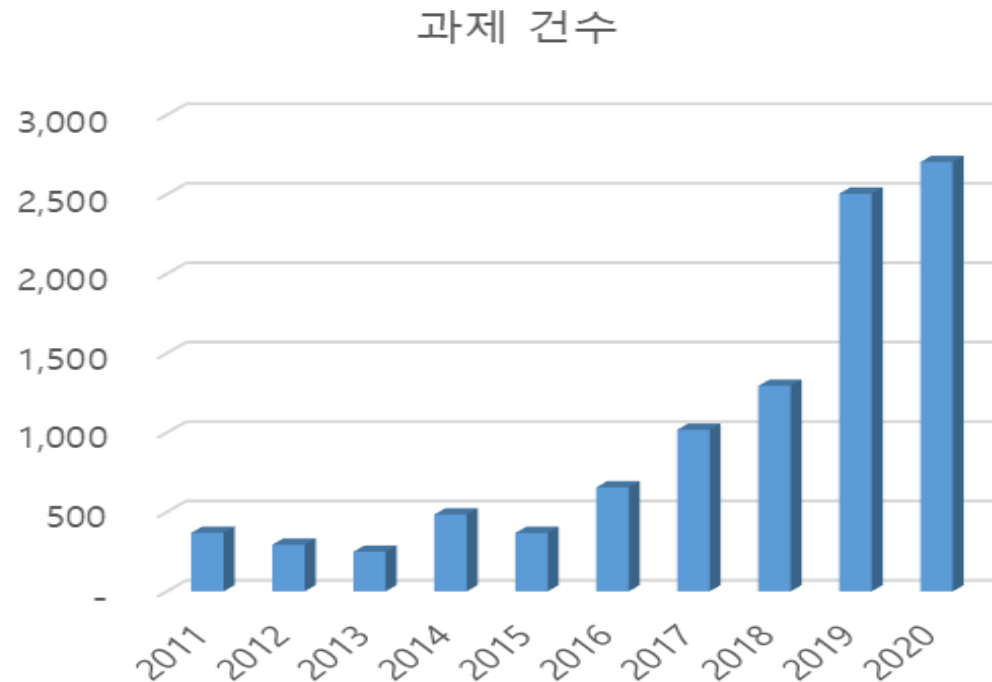
국가R&D 지식정보와 연구데이터를 연계하여 R&D 전주기적으로 서비스를 제공



5. 연구데이터 개방과 활용 활성화 방안 – 인프라(2/2)

☑ 우리나라 인공지능 기술을 활용한 과제 현황

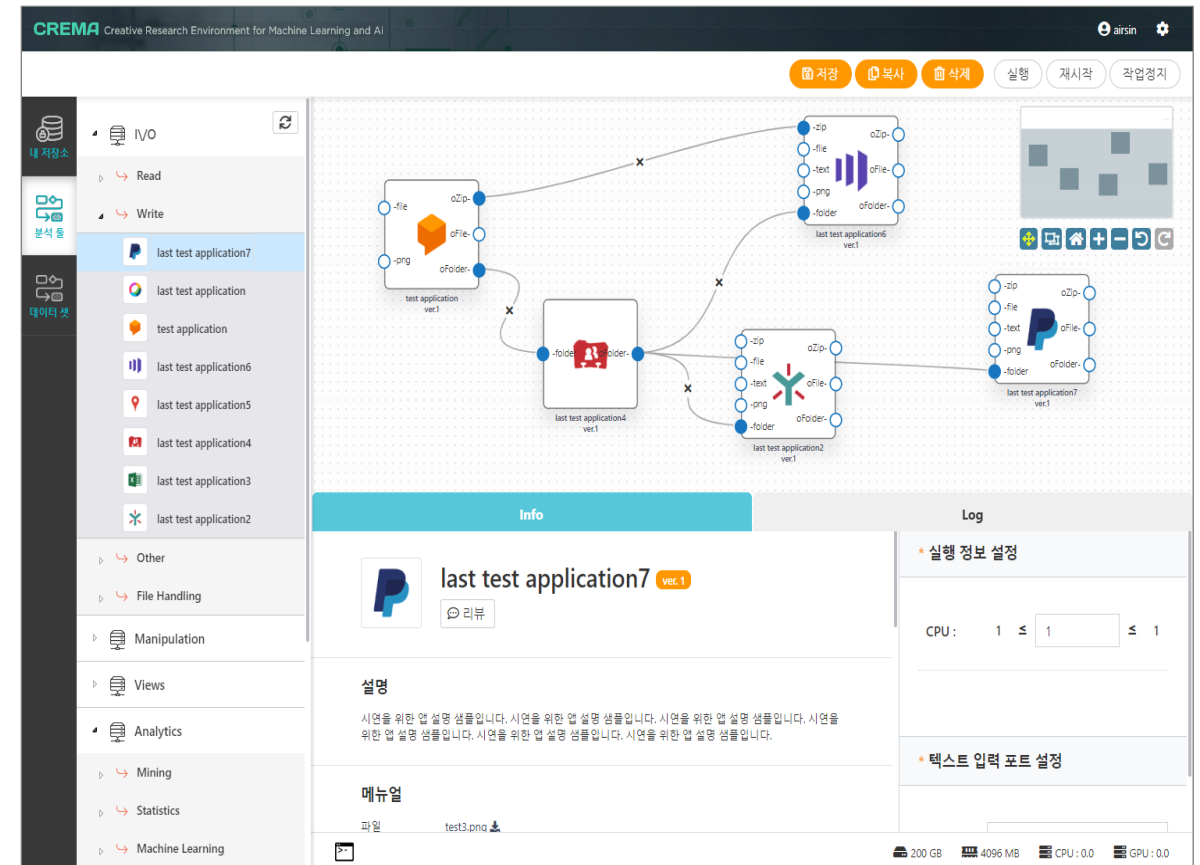
- 2016년 3월 딥마인드의 알파고 발표이후 인공지능에 대한 관심의 증가와 함께 인공지능 기술을 이용한 과제 급증
- 연구자들에게 연구데이터와 인공지능을 활용 가능한 환경 제공 필요



* NTIS 과제 분석 결과('20. 8)

☑ 클라우드 기반의 AI 분석 환경 지원

- 워크플로우 엔진을 통한 다중 사용자 작업 동시 수행 지원
- 클라우드 기반의 인공지능 분석 환경 지원



6. 결론

자발적인 참여와 활용 중심의 연구데이터 생태계 구축으로 국가 R&D 혁신과 효율성 제고

인식의 전환

데이터의 공유가 새로운 가치 창출

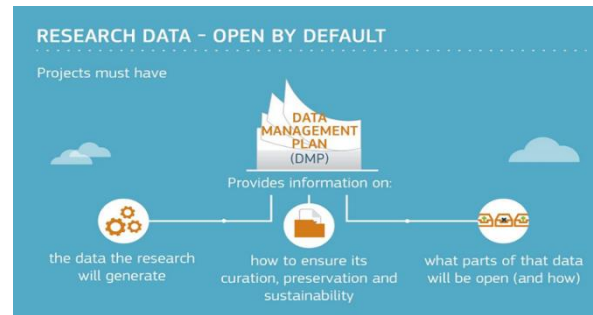
- 국가 공적자금으로 생성된 결과물은 개인의 소유가 아닌 국민의 소유
- 데이터 공유는 손해가 아니라 새로운 과학적 발견으로 가치 있는 지식재산의 근간



DMP 제도 정착

DMP는 공유의 끝이 아닌 시작

- 자신의 성과물을 공개하기 보다는 타인의 성과물 공유를 요구함
- 연구데이터 공유가 활성화 되기 위해서는 DMP 제도를 정착



활용 중심의 거버넌스

활용 중심의 연구데이터 생태계 구축

- 연구데이터 수집은 매우 중요하지만, 이제는 너무 많은 데이터가 생산되는 시대
- 수집과 관리 중심이 아니라 분석과 활용 중심의 데이터 생태계 마련



**QUESTION &
ANSWER**



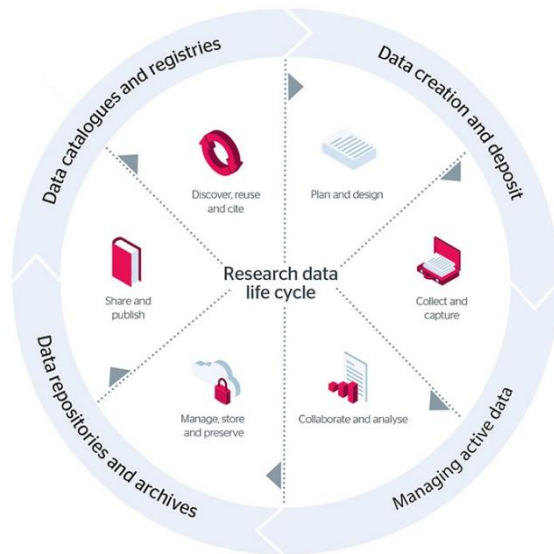
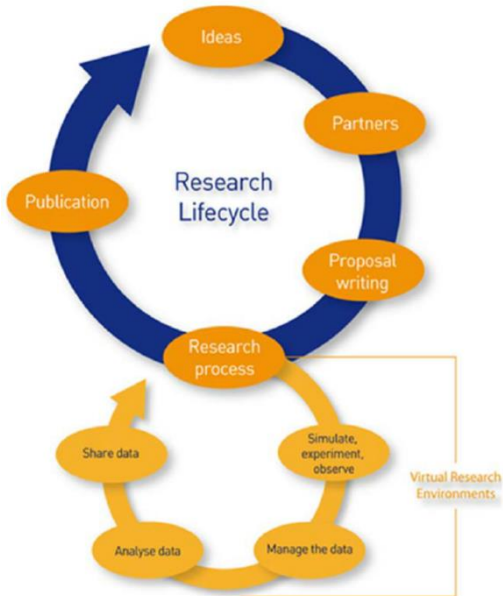
별첨 1 - 연구데이터 정의

국가별 연구데이터 정의는 매우 다양하나 연구개발 활동과 관련한 데이터

☑ 연구데이터 정의

국가 연구개발사업의 각종 실험, 관찰, 조사, 분석 등을 통하여 산출된 자료로 **연구성과의 재현에 필수적이고 객관적인 사실 데이터**
(공동관리규정 제2조)

* 예비분석결과, 논문, 연구노트, 보고서 등은 해당되지 않음



☑ 해외 연구데이터 정의

- (OECD) 수치, 문자, 이미지나 음성 등의 사실적 데이터이며, 주로 과학 연구의 주요한 원천으로 사용하고, 과학자 사이에서 연구결과를 검증하기 위해 필수적인 것
- (미국) 연구 전 주기과정에서 발생하는 데이터
연구계획 단계의 연구데이터는 단순히 연구방법을 보충하는 것이 아닌 연구자가 수집할 정보와 수집 방법, 정보 처리, 분석 계획 등을 포함
- (영국) 연구데이터 관리에 관한 지침서를 통해 데이터 관리차원에서 연구의 전 주기에서 생성되는 데이터
- (호주) 사실, 관찰, 이미지, 컴퓨터 프로그램 결과, 기록, 경험의 형태에서 생성되는 데이터

별첨 2 - 연구데이터 공유의 중요성

연구데이터는 미래 과학과 교육을 위해 사용되거나 재활용될 수 있는 귀중한 자원

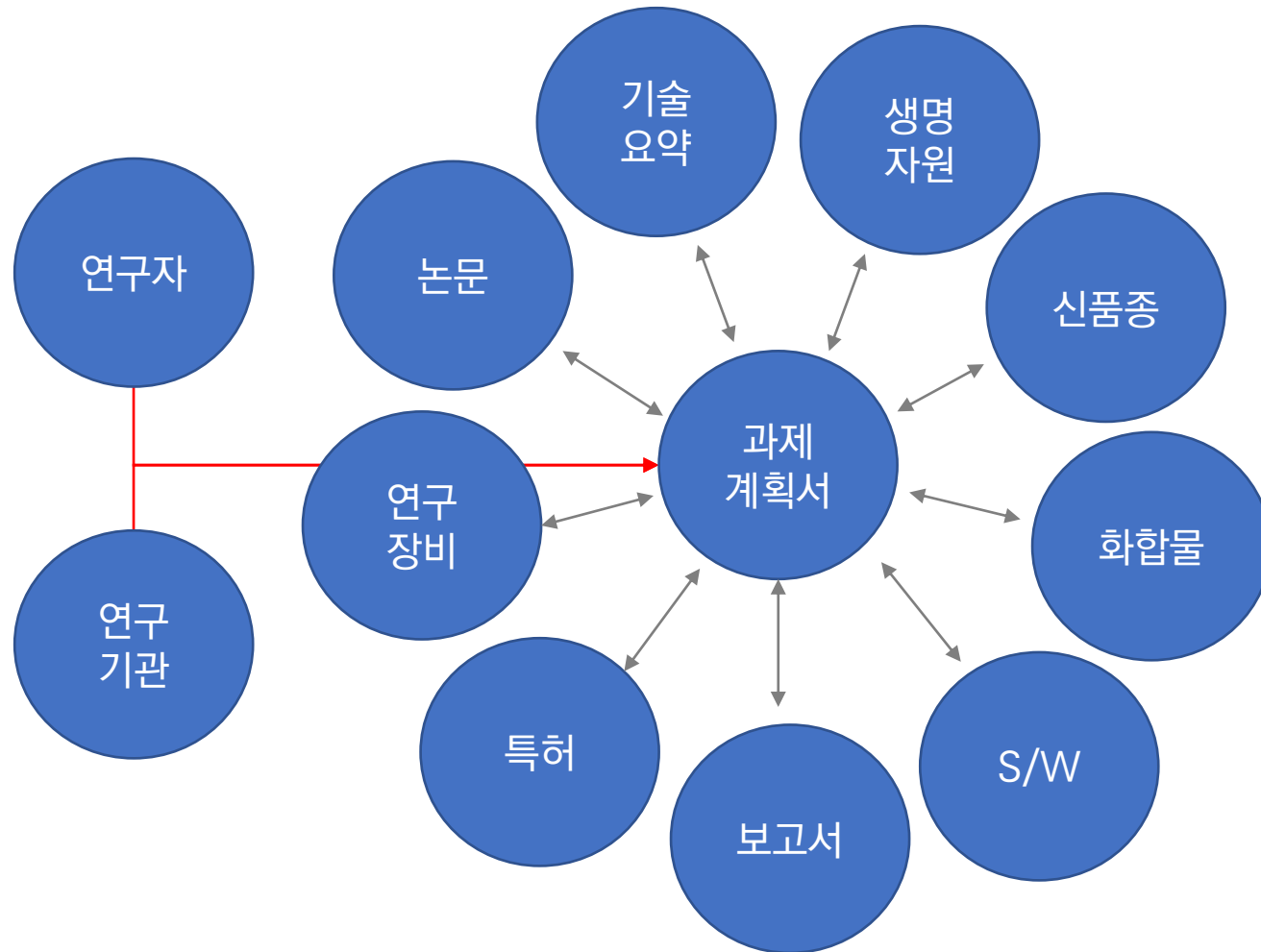


- 과학적 탐구와 토론 촉진
- 연구의 혁신과 잠재적인 새로운 데이터 이용 촉진
- 연구자간(사용자와 생성자)의 새로운 협업을 유도
- 연구결과의 투명성과 책임 극대화
- 연구 결과의 검증
- 연구 방법과 타당성의 확인을 향상
- 데이터 중복 수집 비용 절감(시간, 인력, 노력 등)
- 연구의 영향과 인지도 향상
- 연구자의 직접적인 공헌
- 교육 및 훈련을 위한 중요한 자원

* 참조 : <https://www.waikato.ac.nz/library/services/researchers/managing-information/research-data>

* 참조 : UK Data Archive, 2011

연구데이터를 누가, 어떻게, 어느 과제에서 만들었는지를 알지 못하며, 찾을 수 없는 현실



- 누가 만들었나?
- 어떤 과제에서 만들어 졌나?
- 어디에 보관되어 있나?
- 논문에 사용한 데이터?
- 보고서에 사용한 데이터?

사람 출생 시 출생신고를 하는 것처럼, 연구과제 수행 시 생산되는 데이터 출생신고



[양식 제1호]
출생 신고서
(년 월 일)

※ 신고서 작성 시 뒷면의 작성 방법을 참고하고, 선택항목에는 '영표(O)'로 표시하기 바랍니다.

① 출생자	성명 *한글 (성) / (명) 한자 (성) / (명)	본(한자)	*성별 ♂남 ♀여	*국적 *국혼인중의 출생자 *국혼인외의 출생자
	*출생일시 년 월 일 시 분(출생지 시간: 24시간제)			
	*출생장소 ☐자택 ☑병원 ☑기타			
	부모가 정한 등록기준지			
	*주소	세대주 및 관계		의
	자녀가 복수국적자인 경우 그 사실 및 취득한 외국 국적			
② 부모	부 성명 (한자:) 본(한자)	*주민등록번호	-	
	모 성명 (한자:) 본(한자)	*주민등록번호	-	
	*부의 등록기준지			
	*모의 등록기준지			
	혼인신고서 자녀의 성 본을 모의 성 본으로 하는 협의서를 제출하였습니까? 예☐아니요☐			
	③친생자관계 부존재확인판결 등에 따른 가족관계등록부 폐쇄 후 다시 출생신고하는 경우			
	폐쇄등록부상 특정사항	성명	주민등록번호	-
		등록기준지		
④기타사항	*성명	⑤ 또는 서명	주민등록번호	-
	*자격	☐부 ☑모	☑동거친족	☑기타(자격:)
	주소			
	*전화	이메일		
⑤ 제출인	성명	주민등록번호	-	

- 출생신고는 출생한 날로부터 1개월 이내
- 기간 내에 미신고할 경우 과태료의 제재



- 과제계획서 제출 시 DMP 제출

• 데이터관리계획 (DMP : Data Management Plan)

연구데이터 관리계획(DMP : Data Management Plan)			
① 연구과제요약	사업명		
	과제명		
	주관기관	연구책임자	
	연구개요		
② 연구데이터 형태			
③ 연구데이터 및 메타데이터 표준			
④ 연구데이터 공유 및 제한 계획			
⑤ 연구데이터 재사용 및 배포 계획			
⑥ 연구데이터 보관 계획			

