

SPRI Forum

과학기술연구에서의 소프트웨어 활용 개선 방안

김석원

소프트웨어정책연구소

2014-10-14

소프트웨어 중심사회의 Think Tank



• 경험적 과학

- 자연현상을 직관으로 설명

• 이론적 과학

- 케플러, 뉴턴 이후 모델과 일반화로 설명

• 계산적 과학 (분기)

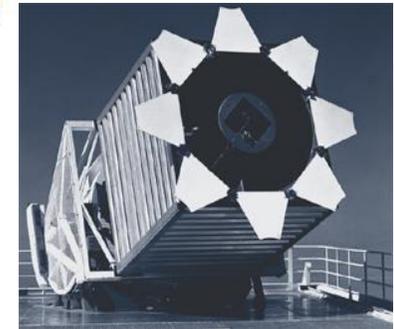
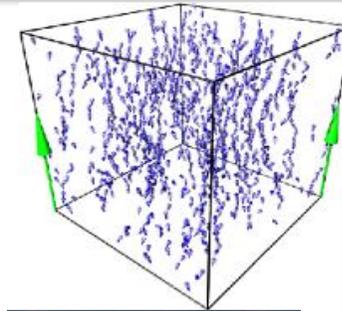
- 이론적 모델이 복잡해지며 해석적으로 풀기가 어려워져서 시뮬레이션으로 검증

• 데이터 중심 연구

- 시뮬레이션 데이터와 실험 데이터의 폭증에 의해 생성되는 대용량 데이터를 소프트웨어로 분석하여 연구



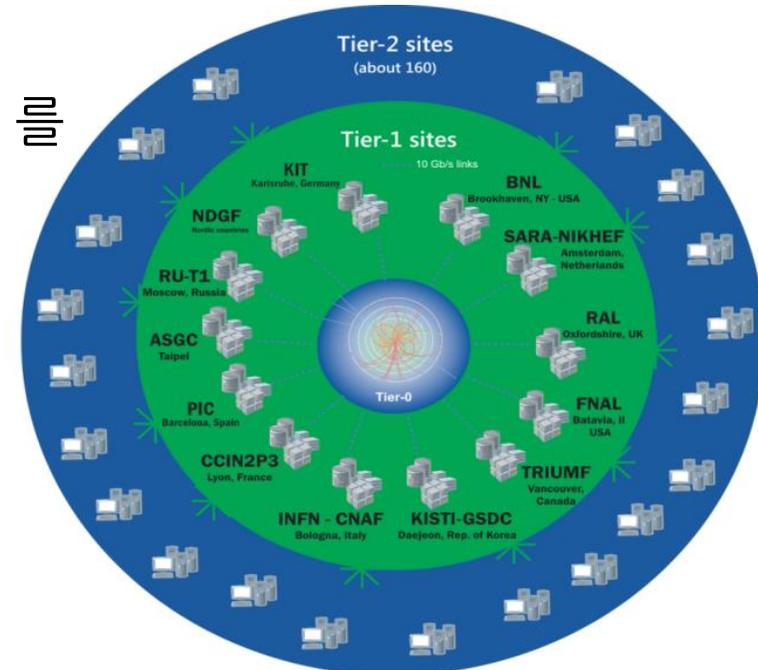
$$\left(\frac{\dot{a}}{a}\right)^2 = \frac{8\pi G}{3}\rho - \frac{kc^2}{a^2}$$



- 2020년의 과학자는 컴퓨터 전문가 수준의 지식 필요
 - 정보 처리의 기본 원리와 기술
 - 시뮬레이션의 효과와 한계에 대한 이해
 - 대형 소프트웨어 시스템의 개발, 시험, 개선에 필요한 개념과 도구
- 컴퓨터과학자는 자연과학자와의 협업을 통해 기술을 적용할 뿐 아니라 새로운 문제를 각성
 - 웹의 개발은 물리학의 필요에서 시작
 - 센서네트워크에서 새로운 네트워크 프로토콜과 알고리즘 개발
- Applied Computing, Computational Science의 중요성이 증가
- 수학과 과학이 서로 협력하여 발전했듯이 과학과 컴퓨팅이 협력하고 영감을 주고받으며 발전하는 시대가 될 것임

Large Hadron Collider

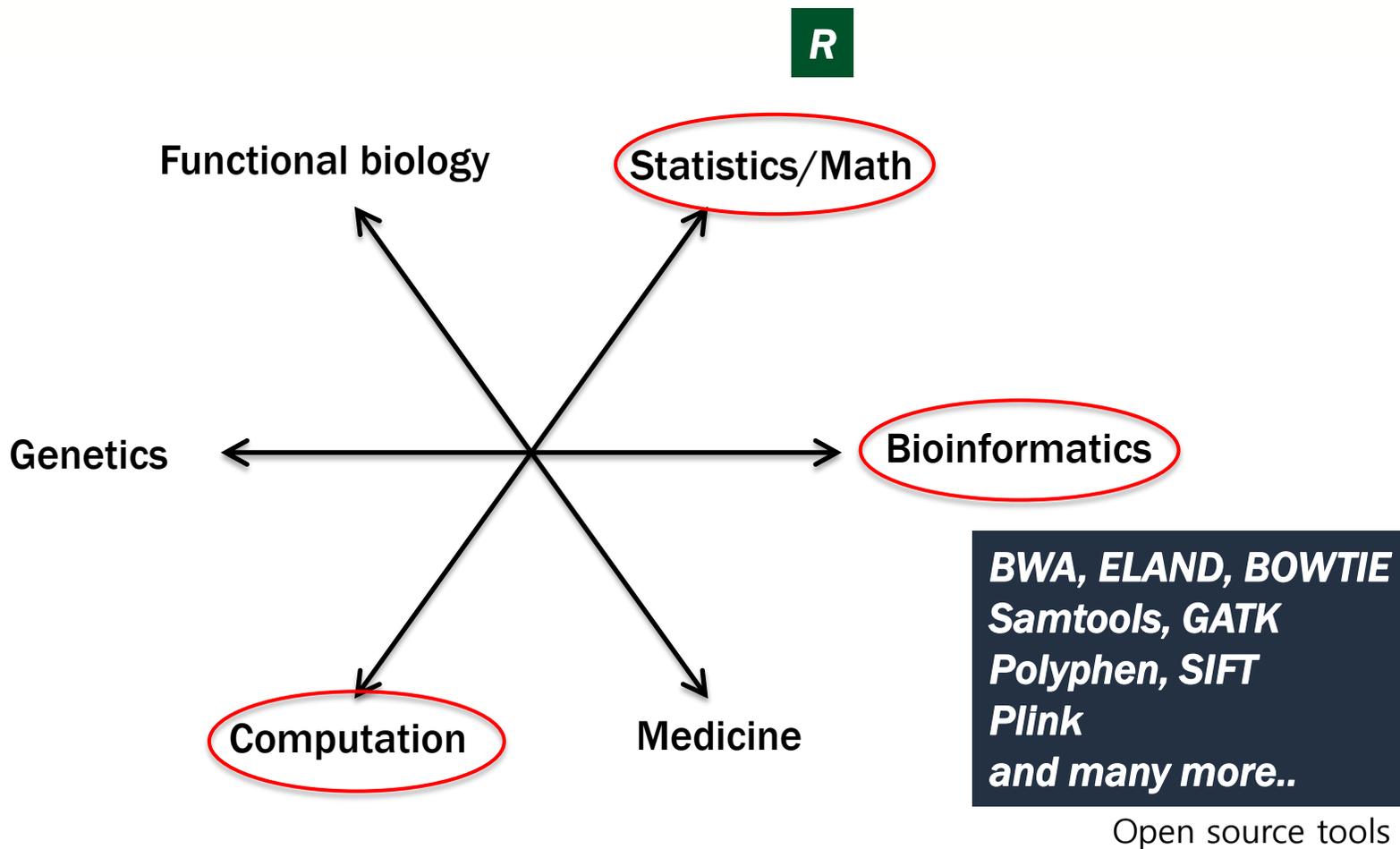
- 양자역학 실험을 위해 스위스와 프랑스에 걸쳐 건설된 대형 강입자 충돌기 (지하 100m, 둘레 27km)
- 매년 30페타바이트의 데이터 생성
- 데이터의 분석, 저장, 관리를 위해 40여개국 170여 기관이 참여하는 Worldwide LHC Computing Grid 를 운영하여 분산 병렬 처리
- KISTI-GSDC (Global Science Data-hub Center)가 2014년에 Tier 1 기관 승인



- 최초의 Human Genome은 13년간('90-'03) 3조원 소요
- 2014년: 3일, \$1000/person
- 1명당 1.5테라바이트 데이터 생성
- 분석이 bottleneck
 - 스트링 얼라인, 매칭, 조합
 - 검색, 시각화, 해석 및 추론
 - 저장, 전송, 압축, 인덱싱, ...



Type	Per sample		
	WGS(70X)	WES(100X)	RNA-seq
Fastq (.gz)	90G	3G	3G
BWA BAM	300G	8G	8G
GATK BAM	300G	8G	8G
Pileup	400G	10G	
Pindel	1.3G	200M	
VarScan	86.2M	20M	
CNV & SV	200M	45M	
Others	300G	8G	8G
Total Size	~1.5 Tb	~40G	~27G

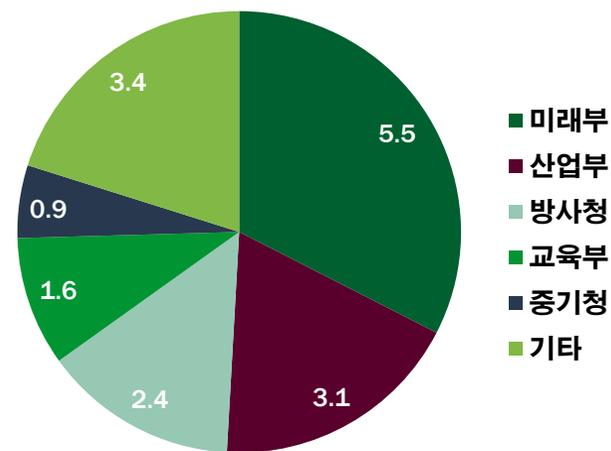
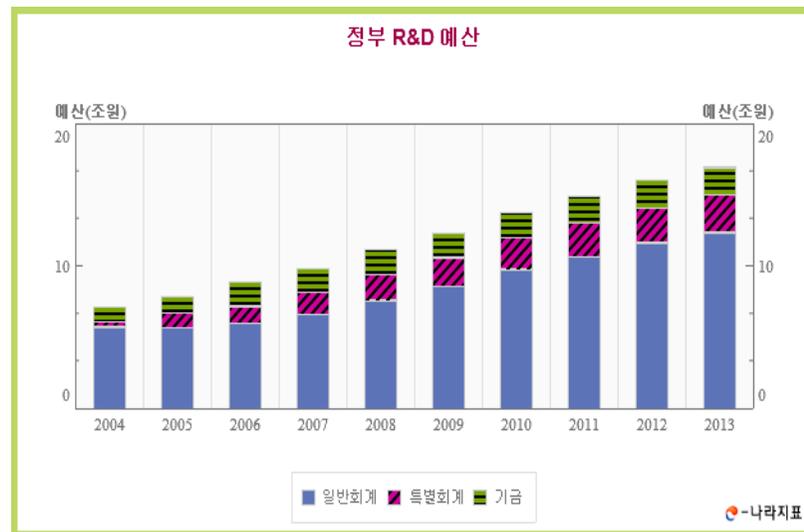


perl, python

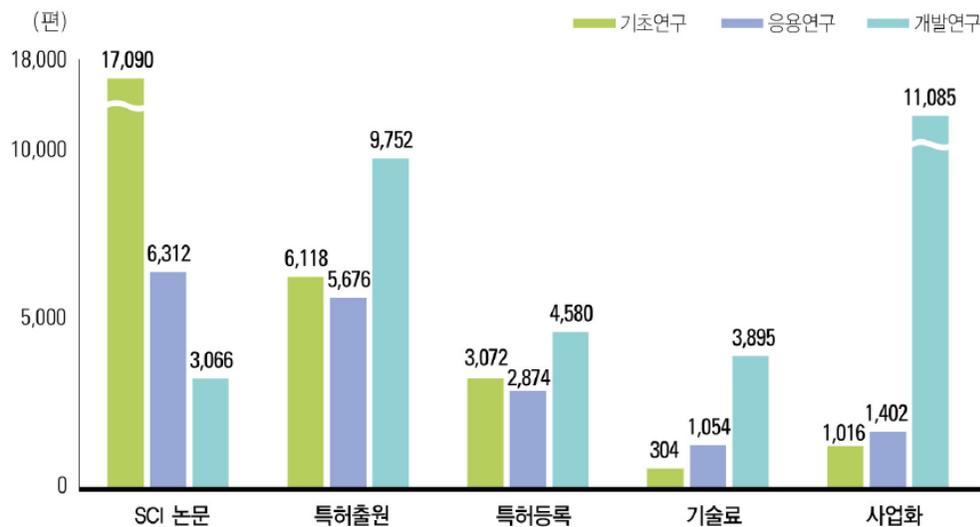
70% dry, 30% wet experiment



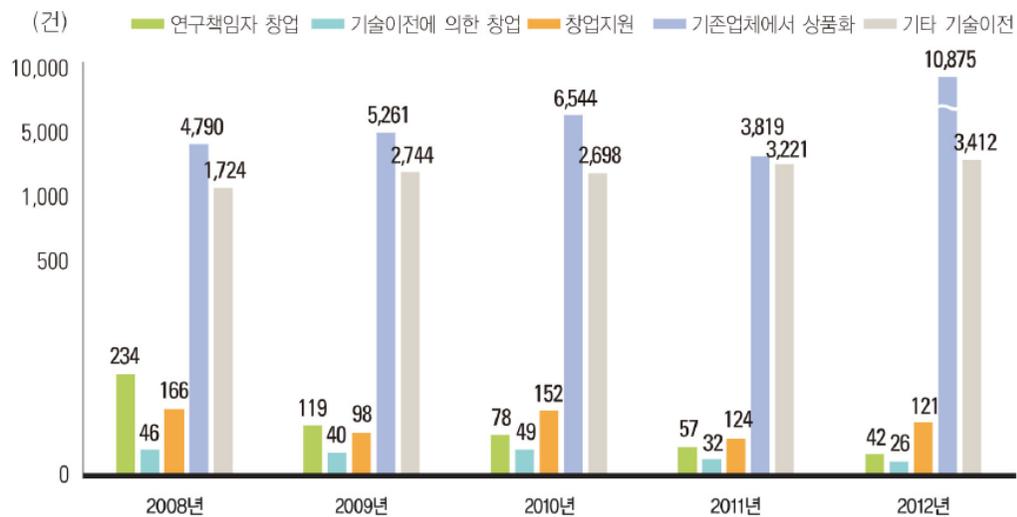
- 2012년
 - ▶ 16조
 - ▶ www.index.go.kr
- 2013년
 - ▶ 16.9조
 - ▶ 2013년도 국가연구개발사업 조사분석 결과 (미래부)
- 2014년
 - ▶ 17.7조 (예정)
 - ▶ 2014년도 정부연구개발 투자 방향 및 관리제도



2013년 부처별 연구개발 투자



2012년 항목별 현황



연도별 사업화

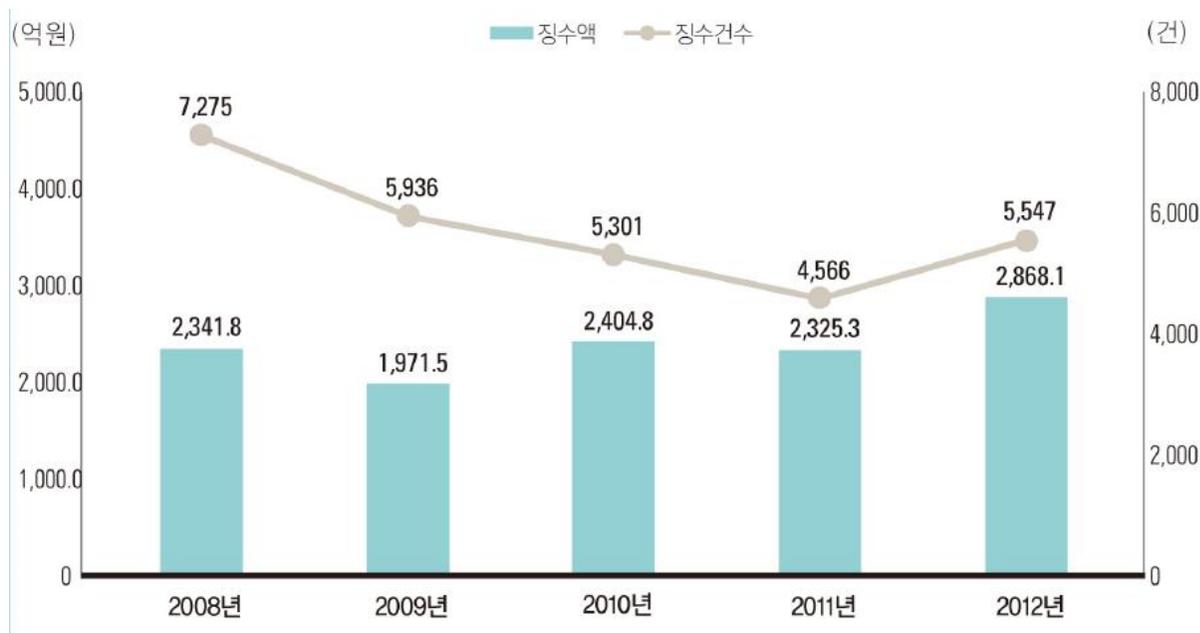
국가과학기술위원회·한국과학기술기획평가원, 「2012년도 국가연구개발사업 성과분석 보고서」, 2013. 12.



	2008	2009	2010	2011	2012
기술료 징수건수	7,275	5,936	5,301	4,566	5,547
기술료 징수액	2,341.8	1,971.5	2,404.8	2,325.3	2,868.1

2012년 총연구비: 16조원

기술료 징수율: 약 1.8%



자료 : 국가과학기술위원회·한국과학기술기획평가원, 「2012년도 국가연구개발사업 성과분석 보고서」, 2013. 12.



- 기술적 성공률 평균 90% 상회, 사업화 성공률 약 20% 수준, 대학, 연구소 4.4%
 - ▶ 선진국 영국 70.7, 미국 69.3, 일본 54.1
 - ▶ 국가연구개발투자의 성과측정방법 연구-국회예산정책처 2013
- 공공연은 19만건 이상의 기술을 보유하고 있으나, 아직까지 15.4만건 이상의 R&D결과물이 사업화되지 못하고 휴면상태('12년)
 - ▶ 10호 제5차 기술이전 및 사업화 촉진계획(안) - 2014. 4.23. 국가과학기술심의회.



- 논문: 학술지, 학술대회, 저역서
- 특허
- 기술확산(기술이전, 기술지도, 기술평가), 사업화, 기술실시계약
- 인력활용 및 양성
- 기타: 국제학술대회 초청강연, 공동/위탁 연구추진

2012년 성과	기술확산			사업화		2012년 계속과제
	기술이전	기술지도	기술평가	사업화완료	사업화추진중	
기초연구사업	2	655	0	7	46	5258
원천기술개발사업	0	61	0	3	45	1322
원자력기술개발사업	0	181	0	0	15	574
거대과학사업	0	11	0	N/A	N/A	141
과학기술국제화	0	115	0	N/A	N/A	367



- 기술료, 사업화의 기여도 낮음
- 현행 성과평과 개선 대책과 병행하여 공유 생태계 활성화 필요
- 소프트웨어와 데이터의 공개를 통한 연구 결과의 재활용 촉진
 - ▶ 공개에 적극적인 연구자를 위한 기반 환경 및 제도 제공
 - ▶ 국가적인 반복 개발 비용 절감
 - ▶ 피어리뷰에 따른 소프트웨어와 데이터의 품질 검증 및 개선
 - ▶ 연구 결과의 빠른 확산에 의한 국가적 연구 역량 제고
- 연구자는 논문, 특허, 저작권 등으로 권익 보호



- NTIS를 통한 과제 정보 공개
- 저작권위원회에 연구산출물 등록
- SW자산뱅크 등록

SOFTWARE ASSET BANK
SOFTWARE ASSET BANK
DETAIL VIEW

기술 표현등급 ★★★★★
관심정보 등록
인쇄하기
다운로드

> 기술 정보

> 제품 정보

▼ 개발 정보

- 개발 목표
- 타겟 고객 및 시장
- 경쟁력
- 세부 구성 기술
- 개발기간 및 소요공수
- 적용한 표준
- 다국어 지원

> 품질 정보

1명이 해당 기술/정보를 추천합니다.

추천하기 >

★★★★★ 사용자평가 6.0

별점주기 >

AUTOSAR 기반 차량제어기 SW플랫폼 및 개발도구 상용화 개발

- 통합 환경 내에서 개발 신 영역을 내증할 수 있는 솔루션을 제공함으로써 신 장제품 개발공정에 있어 노력 및 비용을 절감할 수 있음

R3.1 개발도구 정병력 요소 (핵심)

- 도구 사용 편의성 제공을 위한 Wizard 기능**
 - 사용자의 오류를 줄이기 위한 사용자 유도
 - 사용자의 임의 빈도를 줄일 수 있도록 기본적으로 설정된 Default 값 생성
 - ECU 설정 단계별 설정 편의 최적화
- 통합 도구 체인**
 - 도구업체 최초 시스템/ECU 설정도구 통합
- Defacto 진정 표준데이터 포맷 변환**
 - Non-AUTOSAR 정장 표준파일(DBC, OIL 등)의 AUTOSAR 도구 설정포맷 변환
- 진행상황 Progress 뷰**
 - 시스템 설정 단계
 - ECU 설정 단계
- 직관적인 사용자 인터페이스**
 - 사용자 경험(UX)을 적용한 쉬운 접근방식
 - 일반적인 시스템 레벨의 모델기반 디자인
 - 시스템 설정 집합의 통합 솔루션 제공

구성 기술

- AUTOSAR R3.1 플랫폼 (통신 및 시스템 기능)
- Network Management Interface, Communication Manager

세부구성기술

- 요약보기 >
- 분류체계 상세보기 >
- 연관기술 상세보기 >
- 첨부파일 상세보기 >
- 문의하기 바로가기 >





- 20기관, 42명
- 연구결과의 공개에는 대부분 긍정적
 - ▶ 85% 찬성
 - ▶ 반대의 경우에도 조건부 찬성 의견 (사업화 대상은 제외, 핵심기술 제외, 연구자의 자율권 보장 등)
- 강제적보다는 자발적, 기계적보다는 선택적
- 공개를 위한 오버헤드 우려
- 핵심 노하우의 보호



■ 기본 원칙

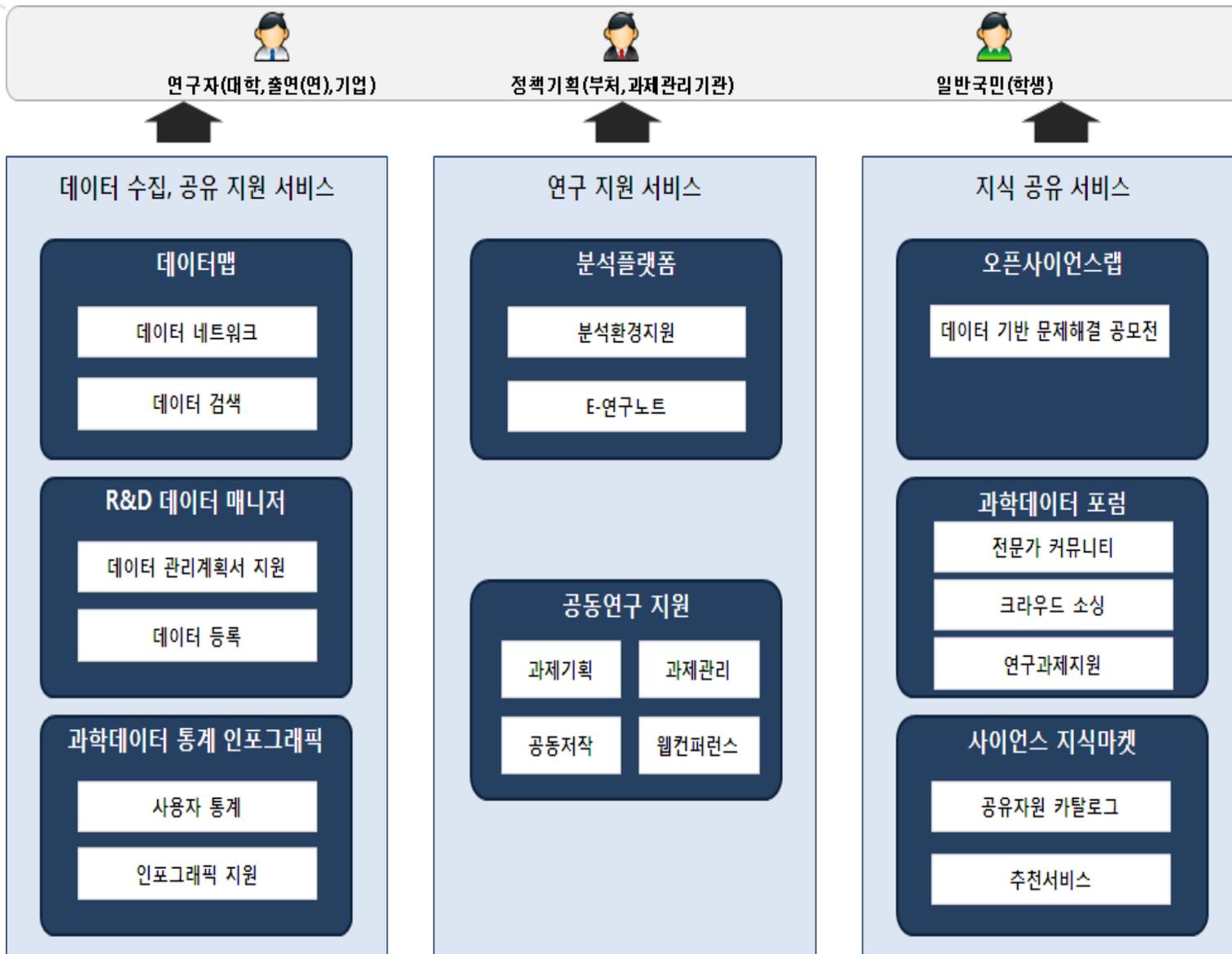
- ▶ 공개를 장려하는 연구 환경 구현이 목적
 - 널리 활용되는 산출물의 가치 인정
 - 공개후의 품질 개선, 공동 개발, 기술 전파 등의 활동에 대한 인정
- ▶ 연구자 판단에 의한 공개 범위 결정
 - 연구 특성에 따른 차이
 - 공개 소프트웨어 내용
- ▶ 사업화에 실패한 산출물의 공개 의무화
 - 급변하는 소프트웨어 환경의 특성상 일단 사업화가 실패하면 같은 코드로 시도하기 어려움
 - 공개를 통해 다른 사람이 활용할 기회를 주는 것이 국가연구개발제도 취지에 합당

■ 공개를 위한 추가 작업에 대한 인정, 지원

- ▶ 연구 계획 수립에 반영하도록 유도



- 과학기술 빅데이터 공동활용 종합계획(안)('13~'17)
 - ▶ 데이터 수집, 공유 지원 서비스
 - ▶ 연구지원 서비스
 - ▶ 지식 공유 서비스
 - ▶ 표준화, 메타데이터, 품질관리
- 과학기술데이터 보존 관리 의무
- 시범사업
 - ▶ 오픈 사이언스 랩
 - ▶ 사이언스 데이터 맵
- NTIS는 메타정보 관리, 각 전담기관에서 데이터 관리





- 공식 산출물
 - ▶ 애플리케이션, 플랫폼, 프레임워크
 - ▶ 공식 산출물의 임베디드 애플리케이션/소프트웨어
- 연구도구 및 부수적 소프트웨어
 - ▶ 시뮬레이션 프로그램/툴, 상용프로그램의 커스텀 프로그램
 - ▶ 데이터 프로세싱 & 유틸리티
 - ▶ 시각화 & UI
 - ▶ 기타 연구개발 과정에서 개발되는 소프트웨어
- 구성
 - ▶ 프로그램 코드, 관련 문서, 샘플 데이터 및 샘플 프로그램
 - ▶ 원칙적으로 제3자가 공개된 내용을 참고하여 빌드하고 사용할 수 있도록 하는 일체의 구성품



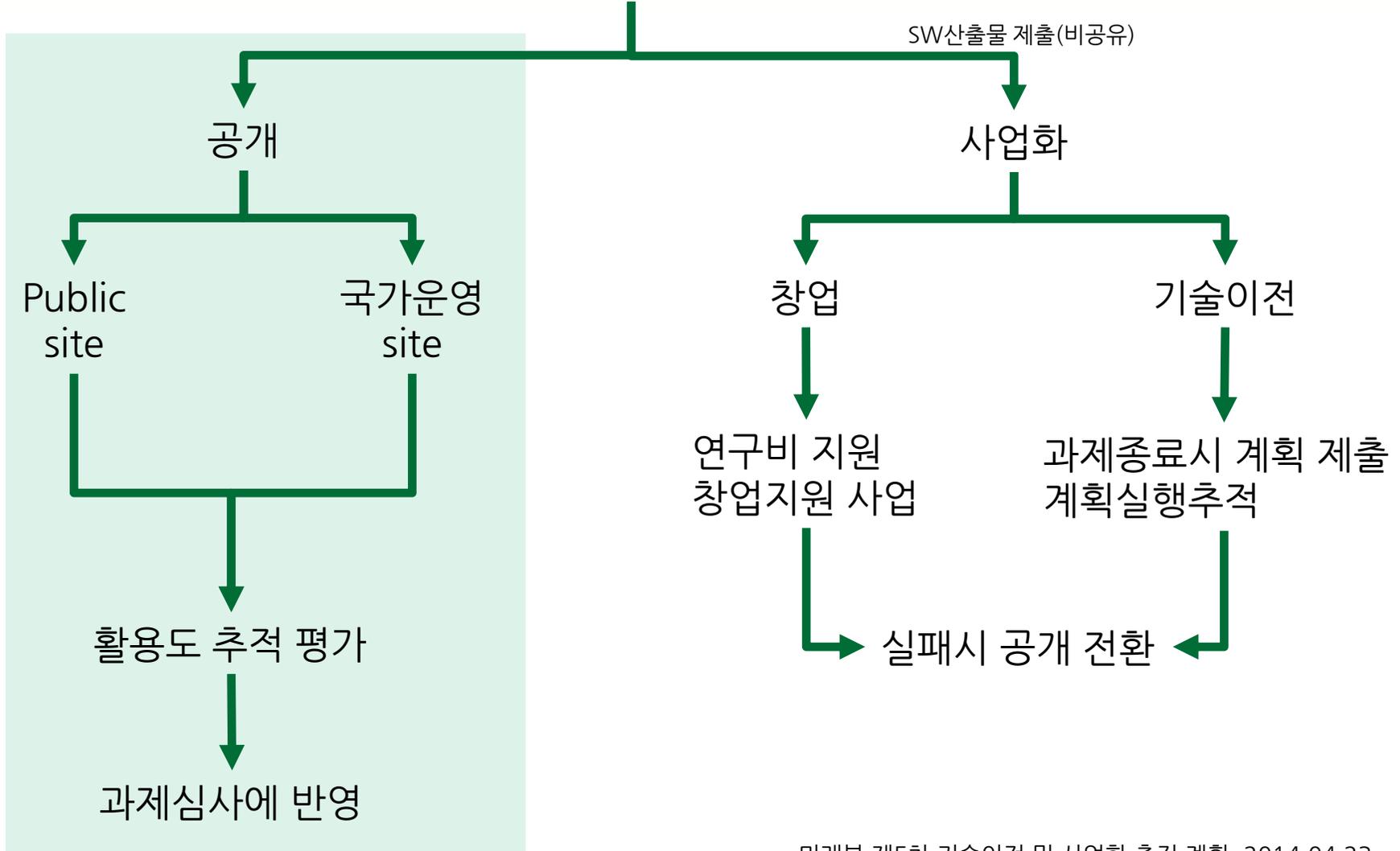
제안, 심사

과제수행 & 종료

평가체제

인센티브

SW산출물 명시, 심사





- 과제 제안시 소프트웨어 산출물 명기
 - ▶ 공식, 비공식 소프트웨어 산출물 명기
 - ▶ 공개를 고려한 연구 계획 수립
- 제안서 심사 단계에서 고려
 - ▶ 공개에 적극적인 연구자의 기여를 심사에 반영
 - ▶ 이전에 공개한 소프트웨어의 평가, 활용도, 완성도 고려



- 일반 공유 사이트 이용
 - ▶ Github, SourceForge 등
 - ▶ 공유 링크 제출
- 국가운영 공개 사이트
 - ▶ 영국의 ccpforge 같은 모델 <http://ccpforge.cse.rl.ac.uk/gf/>
 - 정부 지원에 의해 개발하고 공유 및 커뮤니티 플랫폼으로 활용
 - 사용자 통제 가능
 - ▶ 개발 및 운영을 위한 조직 필요
 - ▶ 사업화를 위한 비공개 소프트웨어도 여기서 관리
 - ▶ 지속적인 개선, 유지보수, 활용 감시 및 분석
 - ▶ 표준 라이선스 제공: 완전 공유와 상용 제한
- 활용도, 개선 활동, 커뮤니티 참여도 등의 데이터 축적



■ 기술이전

- ▶ 과제종료시 SW산출물(비공개), 기술이전 혹은 사업화 계획 제출 의무화
- ▶ 기업체의 commitment까지 시한 적용 (1년)
- ▶ 시한이 지나고 성과가 없으면 공개로 전환
- ▶ 사업화 실적 부진시 공개로 전환

■ 창업

- ▶ 개발 연구원 창업시 경제적, 기술적 지원
- ▶ 초기 연구개발 및 운영자금 지원
 - 연구비 지원 (예. 3년간 1억/년)
 - 현행 창업지원 프로그램 활용 (산업부, 중기청)
- ▶ 유예 기간(예. 3년) 동안 기업 가치를 높여 신규 투자를 유치하거나 일정 규모 이상의 매출 실적을 올리는 것이 목표
- ▶ 전문가 멘토링 제공 (창업지원 프로그램과 연계)



- 모든 과제는 이미 공개한 경우를 제외하고 SW산출물을 과제 종료시 과제 공개사이트에 게시 및 1차 평가
 - ▶ 산출물의 품질, 문서 품질, 샘플 코드 등
 - ▶ 사업화 과제 산출물은 공개 유예
 - ▶ 연구 진행 중 공개하여 활용하는 것을 권장
- 공개
 - ▶ 공개 자료 활용도, 개선을 통한 완성도, 커뮤니티 참여도 기준
- 기술이전
 - ▶ 미달성시 제출된 산출물을 공개로 전환
 - 1년이내 기업체 commit 미달성
 - 기업체가 3년이내 사업화 실패
- 창업
 - ▶ 3년 이내 투자 유치 혹은 사업화 성공
 - ▶ 미달성시 산출물을 공유로 전환



- 공개되는 데이터, 소프트웨어 등의 평가, 관리, 추적
 - ▶ 활용 상황 감시 및 평가
 - ▶ 활용 확산을 위한 인프라 구축 및 운영
 - ▶ 접근성 향상
- 연구개발 및 지원 서비스
 - ▶ 과제의 소프트웨어 개발 지원 - 공개를 위한 표준, 컨설팅, 개발
 - ▶ 공개된 산출물의 활용 확산 - 교육 프로그램, 산출물의 서비스화 등
 - ▶ 자체 연구 개발 - 과학기술연구의 기반구조
 - Computational Science
 - 교육, 공유 서비스
 - 협업 과제 참여를 통한 인력 개발
- 정기 학회, 워크숍, 커뮤니티 운영 등 정보 교환 창구 확대 및 강화
- 기존 관련 조직 활용
 - ▶ KISTI
 - ▶ 소프트웨어공학센터