

기획 착수 회의

2016년 11월 16일

- **The change of scale has led to change of state.**
 - 큰 규모를 활용해 작은 규모에서는 불가능했던 새로운 통찰이나 가치 추출 (예: 스틸 사진의 모음 → 영화)
- **Big Data를 가능케 만든 현대기술**
 - 디지털 기술 → Massive Datafication (예: Digital Universe Project)
 - 인터넷 기술 → Unlimited Accessibility
- **Big Data 시대의 변화**
 - 거대 규모의 데이터를 통째로 분석 (No sampling, No hypothesis)
 - 데이터 정확도에 대해 관대 (Redundancy가 accuracy를 담보)
 - “왜”에서 “어떻게”로 : 인과관계 보다는 상관관계에 집중

많은 연구소들이 위협에 처할 것이다.



포럼의 목표(안) 및 구성

목표 1 : KIST내 data-driven research 활성화를 위한 제안 도출

- Data collection and management 인프라/제도
- Data 활용기술 개발
- Data 공유를 위한 제도적 장치

목표 2 : KIST내 시범사업 제안

기획/기반기술	소재	바이오	농학
강대신 서덕록 오정수 김수현(?)	안재평 이광렬 김찬수	김승종 이득희	최용수

Forum Agenda

- 분야간 빅데이터 현황 파악 (KIST내외부 포괄)
 - KIST내 분야간 격차가 클 것
- Big Data & Machine Learning 분야 전문가 초청 세미나
- Data-driven 연구 기반을 위한 원내 정책제안
 - Data collection & management
 - Data mining and Machine learning team 구성
 - Open Data 정책 / 운영안
- 분야 (주제) 별 data-driven research 과제 제안기획
- 연구사업 추진전략
 - KIST내 시범사업 추진 전략
 - KIST외부 연구진과의 협력체계

- 서론
 - 빅데이터의 개요와 활용 현황 (강대신)
 - Data driven research의 세계적 동향 분석 (이광렬/소재의 예)
 - KIST의 현황 분석 (강대신) 및 이슈정리 (기반기술) (이광렬)
 - 기획의 내용 및 목표 (이광렬)
- 각론 : 분야별 자료 취합 (소재/바이오/농학)
 - 분야별 현황
 - 빅데이터 구축을 위한 인프라 전략
 - Data 관리 및 국제협력 계획
 - Data 활용을 통한 연구개발 전략
- 제언
 - 데이터 축적 및 유지 관리를 위한 인프라 구축 (각론 후)
 - Data scientist의 육성안(강대신)
 - Data 공유를 위한 연구관리 및 평가 시스템 구축 (강대신)
 - Data driven research의 육성을 위한 기관고유사업 운영안 (각론 후)

Tentative Schedule

10/13	빅데이터 머신러닝 관련 전문가 강연 I	2명
11/3	빅데이터 머신러닝 관련 전문가 강연 II	2명
11/16	기획 kick-off	
11/30	분야별 기획내용 발표 및 토론	농학/바이오
12/14	분야별 기획내용 발표 및 토론	소재/기반
12/28	전체 보완 기획 내용 발표 및 토론	
1/11	잔체 보완 기획 내용 발표 및 토론	