

빅데이터와 로봇

Big Data in Robotics

2012.4.12

김홍석 <hskim@kitech.re.kr>

한국생산기술연구원 로봇융합연구그룹

기술의 메가트렌드와 로봇

- ❑ Megatrends: Cloud Computing, Social Network, Big Data
 - 서로 뗄 수 없는 요소 – 기반(인프라/서비스), 연결, 가치
 - 각 요소 별 또는(and/or) 각 요소를 융합한 거대 플랫폼 운용 중
- ❑ 머지않은 장래(10년 내) 자율주행·인식·판단 기술의 출현 예상
 - 예: Google X (무인자동차, 구글 글래스, 로봇 등 100여 가지 혁신기술 프로젝트), Apple Siri...
 - 로봇은 원천기술의 확보가 중요하나 여전히 미흡한 상태
 - 주요 원천기술의 지속적 개발을 위해 (일정 수준의) 개방과 공유는 반드시 필요
- ❑ 현재 IT 거대기업은...
 - 플랫폼 기반으로 사업 전개 (Google, Facebook, Amazon, Twitter, Apple 등)
 - 다양한 개발자를 끌어들이 생태계를 만들고 키우기 위해 리소스 오픈
 - 사용자들로부터 다양한 빅데이터를 수집하여 새로운 가치 창출
- ❑ 그리고... '로봇의 재발견' 작업 필요
 - (꼭 인간을 닮지 않더라도, 인간이 정말 하기 어려운 일을 하는) '도구'로서의 로봇
 - (빅데이터 수집을 위한, 움직이는) '센서'로서의 로봇
 - (노인, 장애인 등) 사회적 약자의 생활을 돕는 로봇

플랫폼은 생태계의 기반: OPRoS*의 기능과 역할

- Robot Software Platform – 기술개발의 틀
 - 지난 5년 간의 기술개발에 매진한 결과, 기술적 완성도는 확보
 - 최근 10개월 여 동안 활용 · 확산을 위해 최선의 노력을 경주
 - Wiki 스타일의 커뮤니티 사이트 구축, 교육·간담회·워크숍·국내외 전시회 참여 등 홍보활동
 - 그러나, 이제 시작인 셈... 지금부터가 더욱 중요!
- Robot Technology Platform – 기술공유와 서비스창출의 장
 - 로봇 요소기술의 공유의 장으로 활용: Repository & App Store
 - 로봇 원천기술의 협업 개발을 위한 장(場) – 클라우드의 활용도 필요: 'Robotics in the Cloud'
 - 기술성과(학 · 연)와 서비스 창출(기업)이 만나는 곳
- Robot Data Platform – 소스코드 뿐 아니라 데이터도 공유
 - 로봇 실험 · 운용 데이터(GB~TB 규모)는 로봇 R&D의 가장 중요한 객체
 - 로봇 데이터는 (소스코드와 마찬가지로) 공유를 통해 가치의 극대화가 가능
 - 연구소는 고가의 로봇센서 보유로 데이터 수집 가능, 대학은 얻기 힘든 데이터를 연구에 활용
 - 로봇 데이터의 축적을 통해 '빅데이터'로서의 가치 창출, 새로운 비즈니스 기회
 - 로봇 자체 뿐 아니라, 향후 (로봇이 활용하는) 센서 네트워크, M2M 데이터까지 포함

* Open Platform for Robotic Services

참고: 제4 패러다임

❑ Science Paradigms*

- Thousands years ago: Science was **empirical**
describing natural phenomena
- Last few hundred years: **Theoretical** branch
using models, generalizations
- Last few decades: **Computational** branch
simulating complex phenomena
- Today: **Data explosion** (eScience)
utility theory, experiment, and simulation
 - Data captured by instruments or generated by simulator
 - Processed by software
 - Information / knowledge stored in computer
 - Scientist analyzes database / files using data management and statistics

Big Data

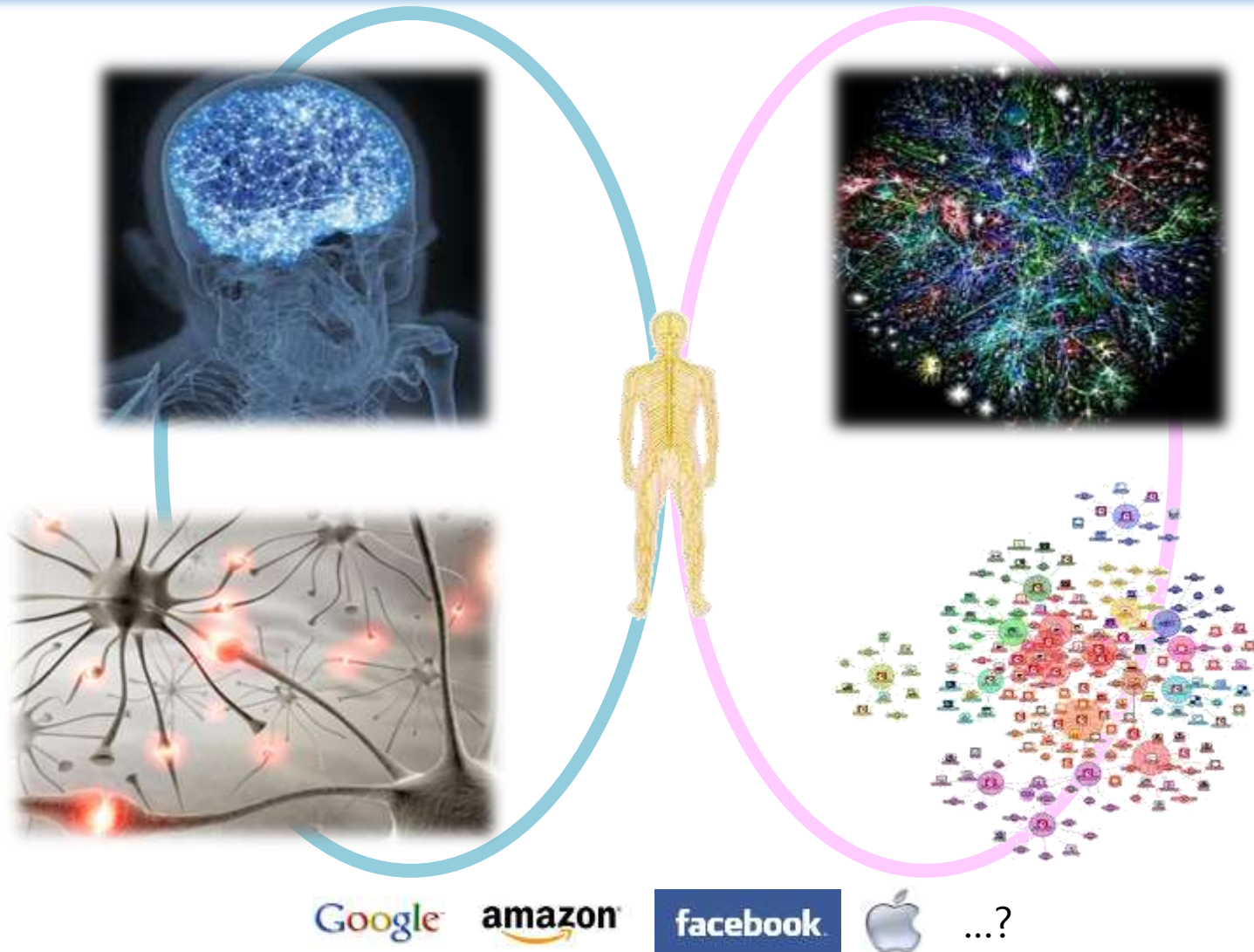
👉 로봇도 이러한 패러다임으로 갈 수 밖에 없음!

* 출처: *The Fourth Paradigm: Data-Intensive Scientific Discovery*, edited by Tony Hey *et al.*, Microsoft Research, 2009,
<http://research.microsoft.com/en-us/collaboration/fourthparadigm/>

최근('11.8월 이후)의 생각...

- 인접 분야(IT)의 기술 트렌드
 - Social Network, Cloud Computing, Big Data...
 - 우리는 과연 '융합'을 하고 있는가?
- 플랫폼 전략, Open Source 전략
 - 우리의 정서와 문화가 문제다
- 로봇의 인식·판단 기술기반이 절대적으로 필요하다!
 - 패턴인식 · 기계학습 겨울강좌 참석('12.2.23~25)
 - 원천기술을 창출·공유할 수 있는 기반과 문화의 필요성
- 그리고, 세대교체
 - 30대~40대 초반이 전면에서 나서야

인터넷 = 뇌



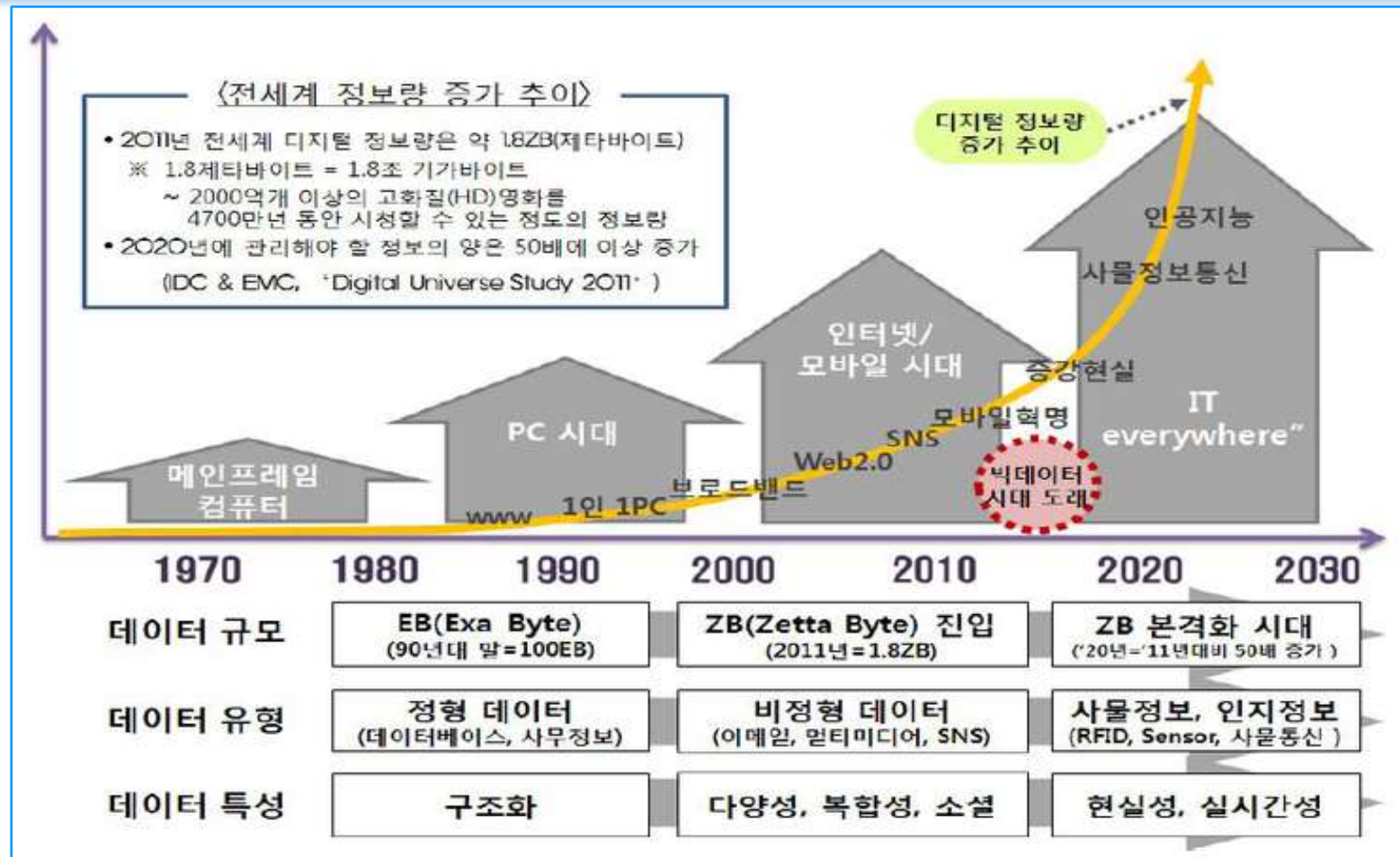
“빅 데이터”란?

- 시스템, 서비스, 조직(회사) 등에서 주어진 비용, 시간 내에 처리할 수 있는 범위를 넘어서는 데이터 [김형준, 그루터¹⁾ 아키텍트]
 - 처리: 단순한 배치(batch)·분석 작업만이 아니라, 실시간 처리 데이터도 포함하고 있는 개념

- Greenplum²⁾의 정의
 - Massive volumes of information – commonly referred to as *big data* – are growing beyond the performance capacity of traditional data warehouse and database management systems.
 - See how big data is challenging businesses to adopt new analytics and architectures to realize business value, and how researchers are preventing disease, combating crime and empowering the developing world.

¹⁾ <http://www.gruter.com/> ²⁾ <http://www.greenplum.com/>

ICT 발전에 따른 데이터 변화



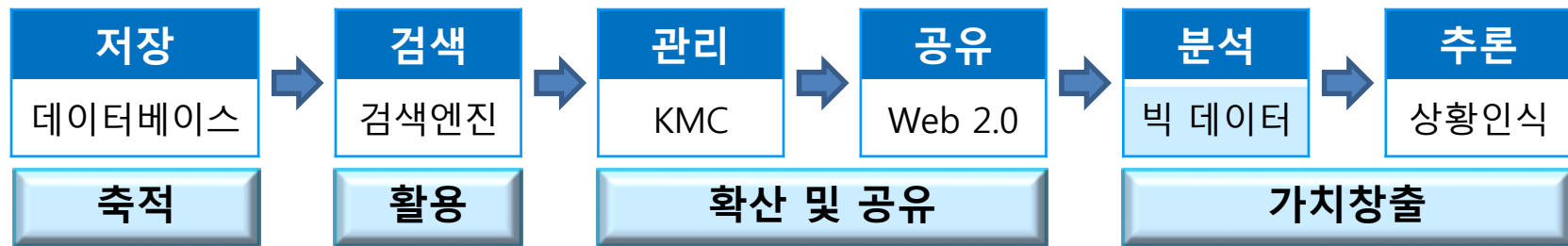
* Kilo (= $2^{10} = 1,024 \approx 10^3$), Mega, Giga, Tera, Peta, Exa, Zetta (= $2^{70} \approx 10^{21}$)

* [출처] 정지선, 新가치창출 엔진, 빅 데이터의 새로운 가능성과 대응 전략, 한국정보화진흥원, IT & Future Strategy 제18호, '11.12.30

빅 데이터는 차세대 이슈

□ ICT 주도권이 데이터로 이동

- 데이터의 변화와 발전은 IT와 비즈니스 영역에서 새로운 공급체계와 수요를 창출하는 원동력으로 작용



□ 시간, 공간, 관계, 세상(그리고 **경험**)을 담는 데이터

- 인적 관계, 위치, 행태, 인식 등 사람과 사회현상을 이해할 수 있는 비정형 대규모 데이터의 증가 및 그에 따른 분석 활용도의 증대

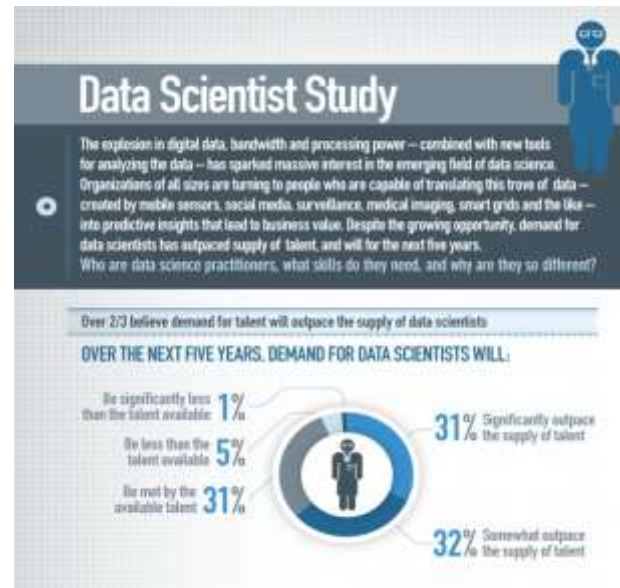
□ 미래 경쟁력과 가치창출의 원천

- 대규모 데이터에서 의미를 찾고 정보와 지식을 만들어 내는 능력이 경쟁력

Data Scientist

□ 데이터 분석 능력과 프로그래밍 기술을 겸비

- 데이터를 분석을 통해 패턴을 찾아내고 추천, 추정, 예측 모델링을 함으로써 다양한 도메인에 적용될 수 있는 기법을 활용할 수 있는 역량을 갖춘 사람
- 기본적인 분석 툴은 통계학, 데이터 마이닝(Data Mining), 기계학습(Machine Learning) 등
- 업무 예: 추천 엔진(인터넷 비즈니스), 스마트 그리드(전력), 각종 센서에서 발생하는 데이터를 수집하여 수율예측 및 문제발생 예방 등의 업무(생산) 등



* <http://mashable.com/2012/01/13/career-of-the-future-data-scientist-infographic/>

빅 데이터 분석 활용사례 및 동향 (1)

□ 구글 독감 트렌드 서비스¹⁾

- 인터넷에서의 집단행동(독감 증상에 대한 검색 빈도)과 의학의 결합
 - 구글은 미국질병통제예방센터의 애틀란타 주 독감 발생 공포에 2주 앞서 예보('08.2월)

□ 오바마 정부의 '노후차량보상 프로그램'과 구글의 예측

- 미 정부는 4개월 간 \$1B 예산 편성, 1주일 만에 예산소진, \$2B 추가 편성('09.7월)
 - 구글은 동 프로그램의 검색 폭주를 보고 호응도를 분석, 예산이 부족할 것을 정확히 예측

□ 구글 트렌드 서비스²⁾

- 구글은 축적된 검색 및 인덱싱 자료를 통해 웹사이트나 키워드의 트래픽 성향을 비교할 수 있는 서비스 제공

□ 미국 CIA는 빅 데이터 해석과 예측분석기술 확보에 주력

- 미국 CIA와 구글은 예측분석기술 전문회사인 Recorded Future에 투자('10년)
 - 웹 상에서 벌어지는 일(개인, 그룹, 이벤트 등)들을 모니터하는 기술

□ 멤피스시(미국 TN) 실시간 범죄감시센터, 사건패턴 수집, 범죄예방에 활용

- 범죄율 30% 감소, 우범지역과 시기를 예측하여 적시적소 인력배치로 범죄 예방

¹⁾ <http://www.google.org/flutrends/> ²⁾ <http://www.google.com/trends/>

빅 데이터 분석 활용사례 및 동향 (2)

- US Military Bets on 'Big Data' to Win Wars¹⁾
 - 미군은 'Big Data' 활용을 위해 매년 \$250M 사용할 것
 - That means creating computer systems that combine “computer speed, computer precision and human agility” to collect and make sense of military intelligence at speeds 100 times faster than today — a power that could also enable military analysts to speedily dig up information from texts in any language. The effort could even lead to truly *autonomous war robots*²⁾ capable of making their own decisions on tomorrow's battlefields.

¹⁾ <http://www.innovationnewsdaily.com/1004-military-bets-big-data.html>, ('12.3.29)

²⁾ Military Battles Information Overload from Robot Swarms,
<http://www.innovationnewsdaily.com/553-battlefield-drones-information-overload.html>

스마트 세상의 불편한 진실

- 사람들은 스마트 기기를 통해 연결된 편리한 생활 구가
 - 폭발적으로 증가하는 규모
 - 전세계 스마트 폰 사용자 5억 명('11년)
 - 모바일 인터넷 가입자 50억 명 돌파('10.7월)
 - 500억 개의 기기가 네트워크에 연결(~'20년)
 - 빅 데이터 기업의 앱과 콘텐츠를 일상적으로 사용

- 빅데이터 기업은 수십 억 개의 '모바일 센서'를 깔아놓은 셈...
 - 사람들은 수 백 달러를 지불하면서도 즐겁게 이 '센서'를 구입
 - 수집되는 빅데이터는 (다시) 새로운 콘텐츠 창출의 기반

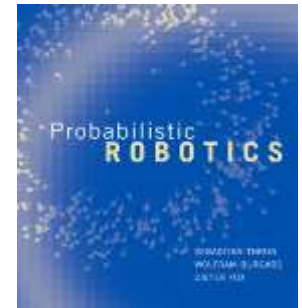
- 빅데이터 시대, '로봇'을 어떻게 볼 것인가?
 - 로봇 자체로 돈을 벌 것인가, 아니면...?

'빅데이터 시대'의 로봇...

- 로봇 알고리즘에서 다루는 데이터는 원천적으로 '빅데이터'
 - Navigation, SLAM, Object Recognition...
 - S. Thrun, *Probabilistic Robotics*, MIT Press, 2005

- R&D의 패러다임 변화
 - (로봇의) 빅데이터는 '오픈 소스'
 - 공유와 협업

- 새로운 비즈니스의 기회
 - 로봇을 어떻게 '빅데이터'의 센서로 활용할 것인가?
 - 또 다른 생각은?



사례: 로봇 빅 데이터

□ Liquid Robotics¹⁾사의 'Wave Glider²⁾'

- Surfboard looking maritime robot
- 하와이에서 캘리포니아까지 2,500마일, 1년 이상의 항해를 통해 데이터 수집
- PacX³⁾ (the Pacific Crossing): the next step... with *Google Earth* and *Virgin Oceanic*. On November 17 Liquid Robotics launched 4 Wave Gliders that will travel across the Pacific Ocean. The four Wave Gliders are expected to collect approximately *2.25 million discrete data points*, and take more than 300 days to complete their voyage... Liquid Robotics is making *this data available to anyone who registers*.



☞ 로봇이 '자연정보수집자' Natural information gatherers' 라는 관점

- 최근 VC의 최첨단 투자 트렌드

¹⁾ 1329 Moffett Park Dr., Sunnyvale, CA 94089; <http://liquidr.com/>

²⁾ <http://liquidr.com/technology/wave-glider-concept/> ³⁾ <http://liquidr.com/pacx/>

참고: 미국 컴퓨터과학 분야 R&D

□ NSF, \$10M/5년 × 4개 과제 지원*

- **3-D Robotics**

- printable programmable machines

- **Socially Assistive Robots**

- 사회적/인지적 결핍자(deficit) 포함 어린이의 사회적, 정서적, 인지적 성장을 촉진(encourage) 하는 로봇의 설계, 구현, 평가를 할 수 있는 computational technique를 개발

- **Making Sense at Scale with Algorithms, Machines and People**

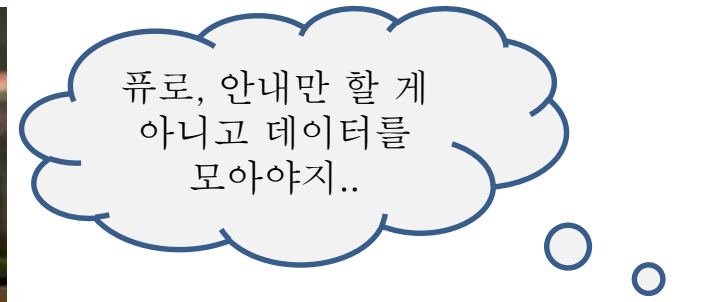
- to investigate tools to make better sense of the volume, diversity and complexity of data being generated by computers, sensors and scientific instruments; this investigation addresses what has become known as the *big data* problem

- **ExCAPE** (Expeditions in Computer Augmented Program Engineering)

- developing an environment where a programmer and an “automated program synthesis tool” collaborate to generate software that meets customer specifications

* <http://sciencebusiness.technewslit.com/?p=8896> ('12.4.3)

뭘 하고 있는 걸까요...?



* 퓨처로봇의 'Furo' (COEX Mall 입구, '12.3.13)

향후 바람직한 추진방향은?

□ 차기 OPRoS 과제에서는...

- 최소한의 선도적인 원천기술 개발은 필요
 - 예: 클라우드 기반 로봇플랫폼 기술
- 활용 · 확산을 위해 비-R&D 성격의 사업 추진 필요
 - 향후 추진과제의 내용은, 홍보, 교육 및 교육 콘텐츠 제작, 로봇 HW 배포, 커뮤니티 운영, 문서화 작업 등 R&D 사업으로는 추진할 수 없는 요소가 대부분

□ 기술 외적인 요소는?

- 소스코드 및 로봇 데이터 오픈 등 개방적 마인드
- 서로 존중하고 배려하는 문화와 협업
- ...

* 작성: 김홍석/2012-04-11, 05-21

* [SlideShare](https://www.slideshare.net/secret/4C2M0dovgthWlG) uploaded 2015-10-13, <https://www.slideshare.net/secret/4C2M0dovgthWlG>

[미래모임 패널토의]

인간과 로봇이 공존하는 한국의 미래사회

2012.5.22 / 삼성호텔

김홍석 <hskim@kitech.re.kr>
한국생산기술연구원 로봇융합연구그룹