

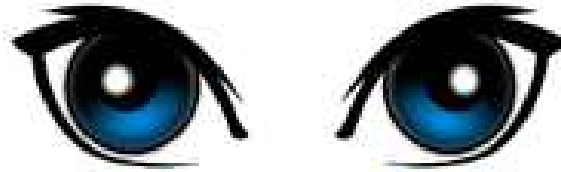
제30회 Open Technet, 빅데이터 오픈소스 플랫폼 기술세미나
2012. 7. 26 목요일

빅데이터 활용 가치 및 국내외 적용 사례

이현종 Leesult 대표
빅데이터 커뮤니티 운영자
<http://www.bicdata.com>

빅데이터를 바라보는 두 개의 시선 – Two Track 접근

기술



분석

Technic 관점	관점	Data 관점
Big data Platform	개념	Big data Analysis
Big data Engineer	전문 인력	Big data Scientist
벤더의 영업 담당자	실무자	조직의 분석 기획자
이공계, 이과	성향	인문계, 문과
과학기술 및 정보통신 연구기관	공공기관	국가 및 공공 데이터 관리기구
하둡, 하둡생태계, NoSQL	오픈소스	Pig/Hive, R
빅데이터 하둡개발자 그룹스터디	그룹스터디	빅데이터 분석전문가 그룹스터디
빅데이터 오픈소스 플랫폼 기술세미나	세미나	빅데이터 분석전문가 오픈세미나
개발자에게서 분석 가치를 찾으려 말라	맹점	기술은 어떻게든 해결되므로 분석만 중요

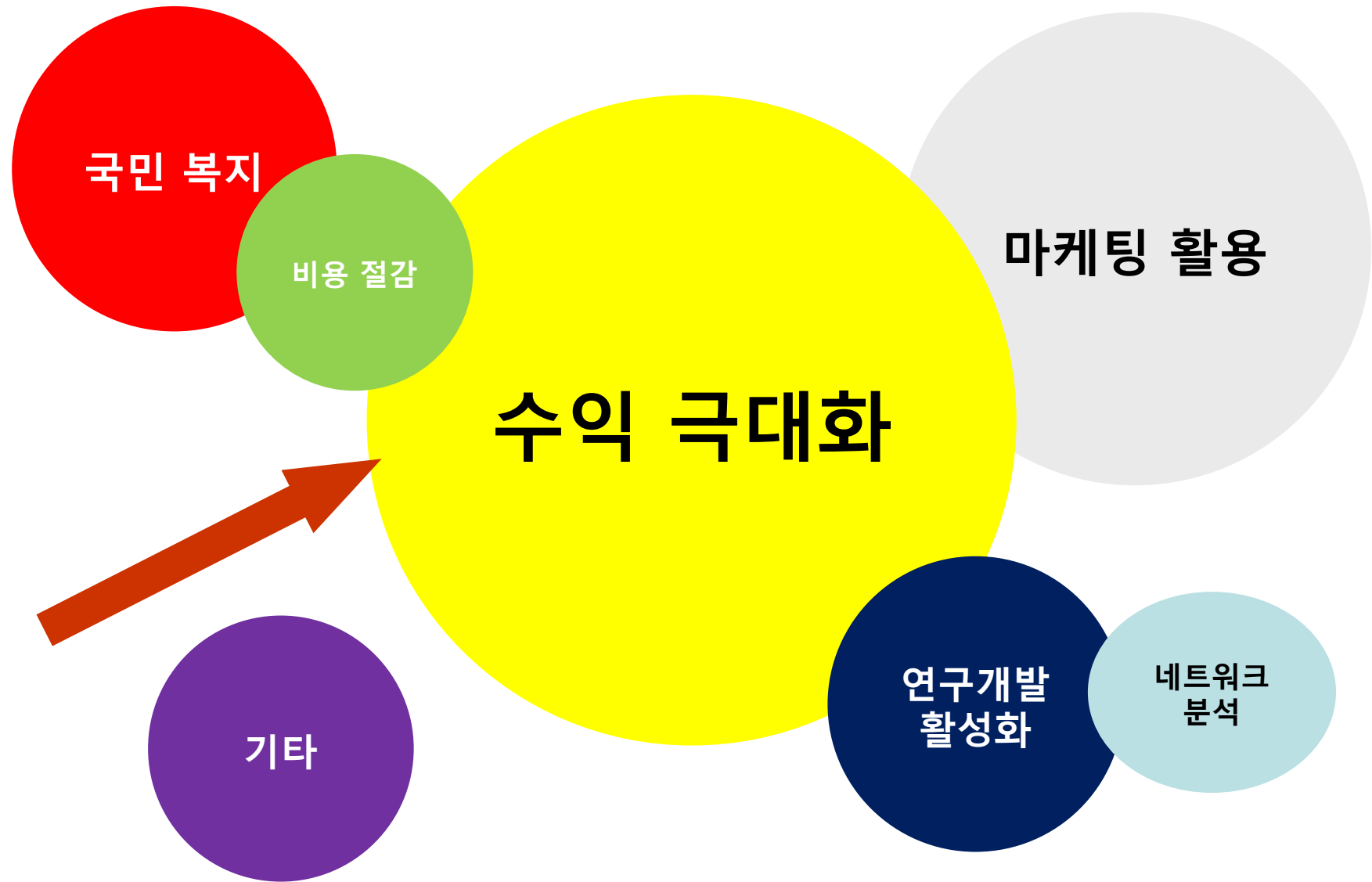
“빅데이터는 기술로부터 잉태되었다”

구글 검색엔진, Big Table, MapReduce
하둡, 하둡생태계
NoSQL, R, 그리고 오픈소스 플랫폼

“기술없는 빅데이터 분석, 실패의 지름길”

(빅데이터 커뮤니티, 이현중, http://www.bicdata.com/bbs/board.php?bo_table=community_free&wr_id=92, 2012. 4)

빅데이터 활용 가치



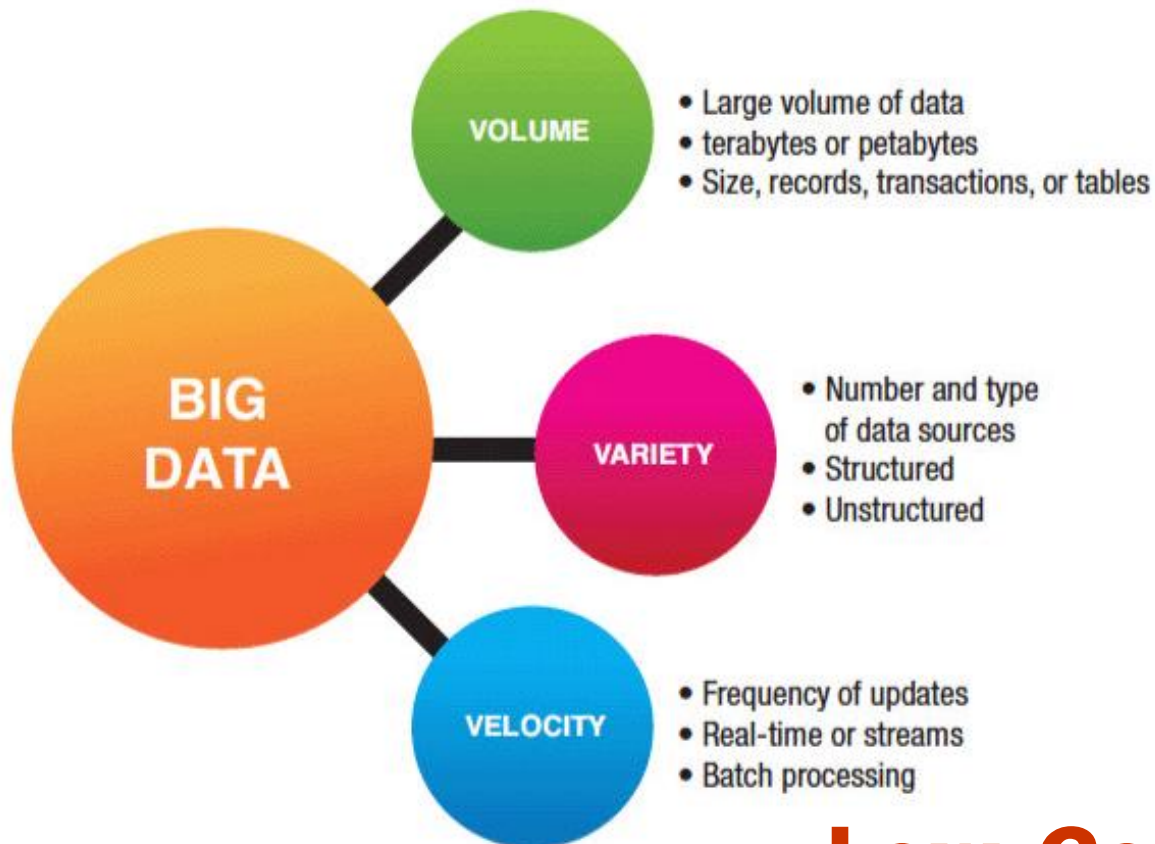
빅데이터 활용 좋은 예

빅데이터 활용, 어떻게 해야 올바른 것일까?



다양하고, 많은 양의 데이터를
저렴한 비용으로
빠르고 원활하게 처리하여
분석 **가치**를 창출

The Three Vs of Big Data



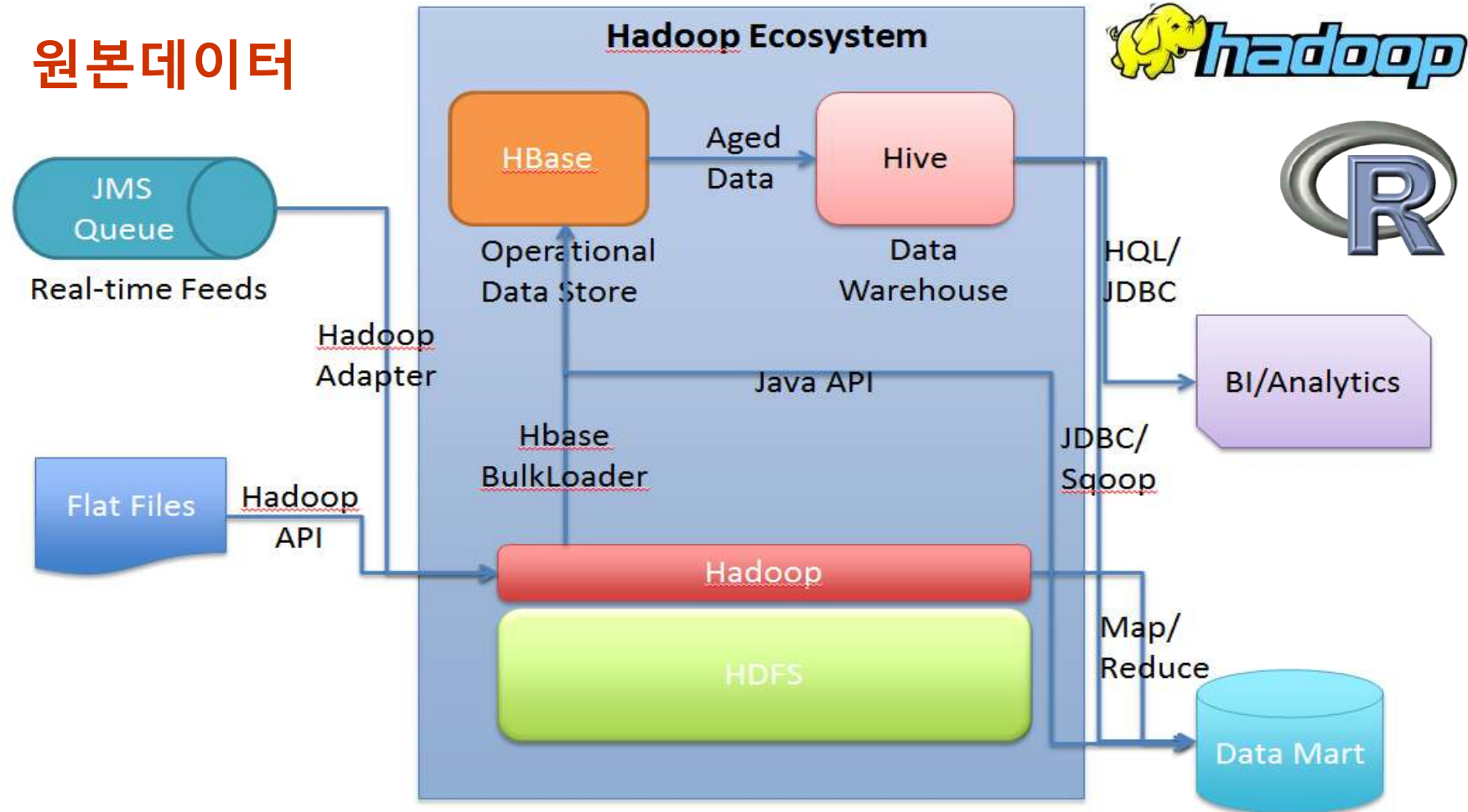
VALUE



Low Cost Platform

빅데이터 오픈소스 플랫폼

원본데이터



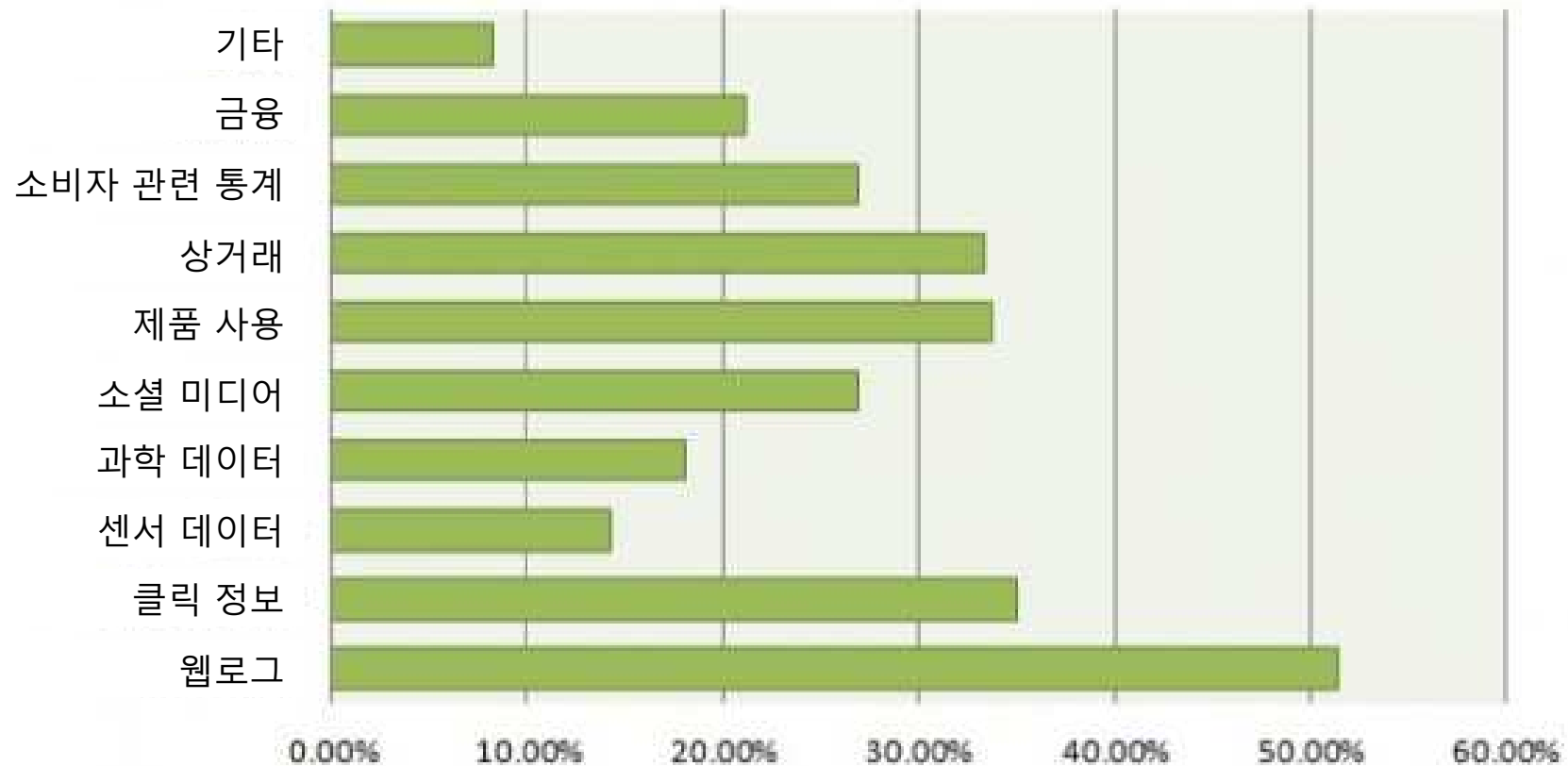
수집 - 저장 - 관리 - 처리 - 분석 - 표현

하둡 활용 데이터 분야

하둡 전문 기업 Karmasphere에서 376명의 데이터 전문가와 인터뷰

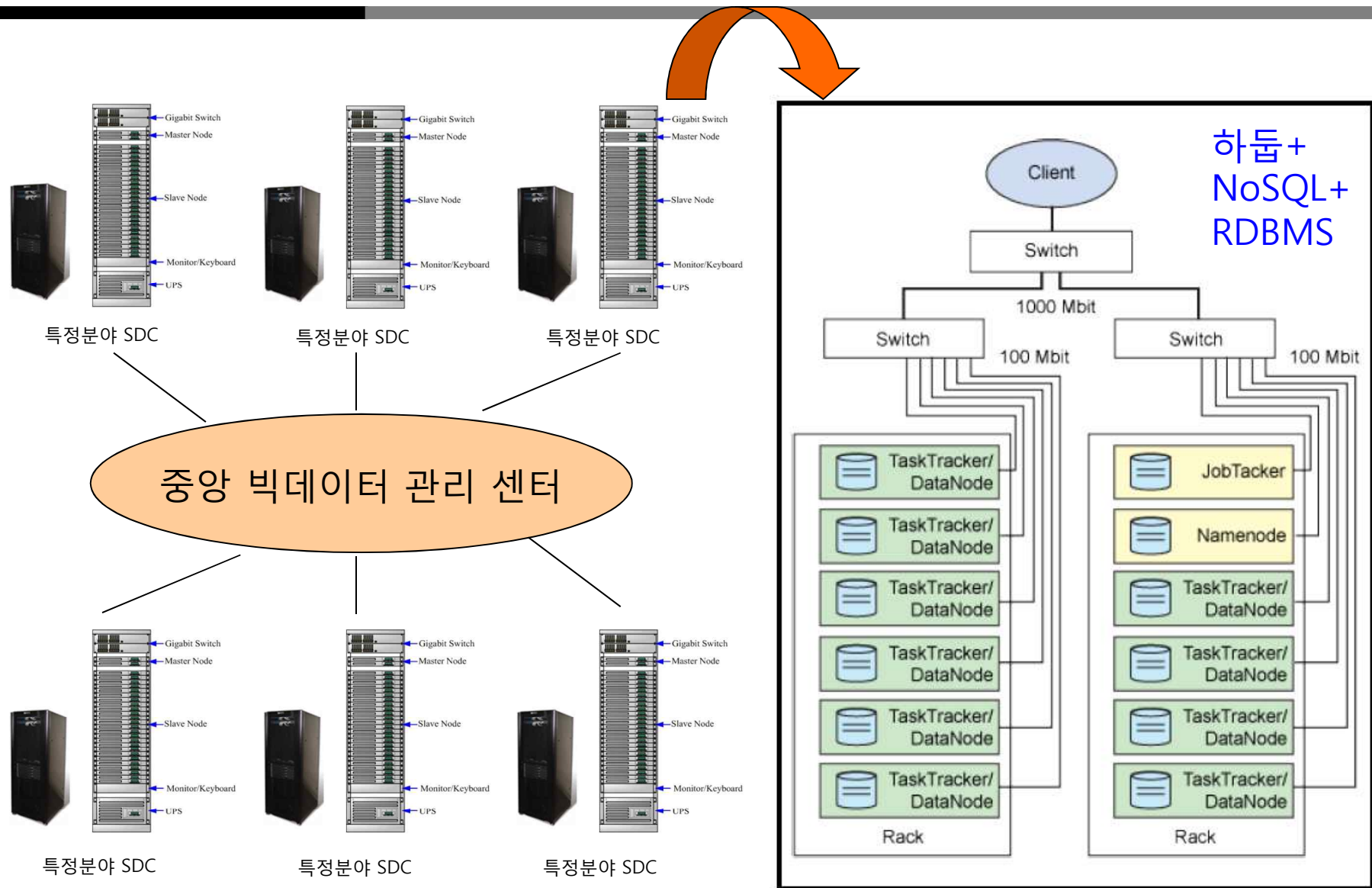
-> 52%만이 하둡 사용중

What kind of data are you analyzing with Hadoop? (Select up to 4)



(기가웜, <http://gigaom.com/cloud/survey-hadoop-clusters-not-that-big-not-changing-the-world-yet/>, 2012. 6)

빅데이터 인프라 H/W, N/W, S/W 구성



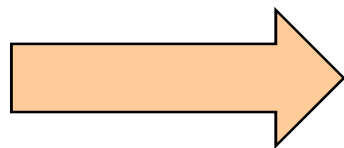
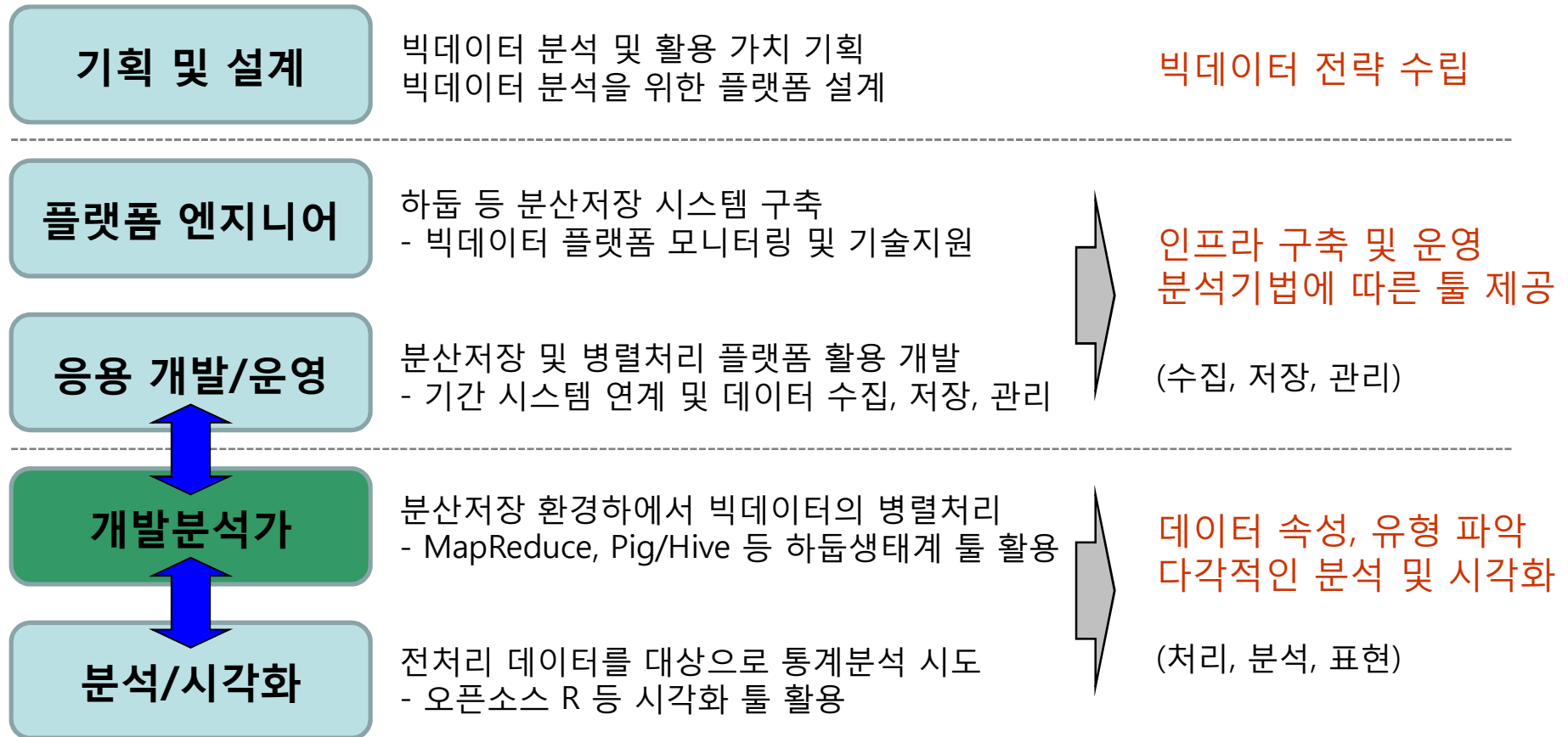
빅데이터도 결국 사람의 일인지라...



빅데이터, 당신의 포지션?



빅데이터 분석 조직 - 5단계 역할



본인 역할 인식과 조직 내 유기적 연계가 관건

빅데이터 국내외 적용 사례

I 빅데이터로 꿈꾸는 행복한 사회

[해외편]

- 1 미국 국세청, 탈세 방지 시스템 통한 국가 재정 강화 • 8
- 2 일본, 센서데이터를 활용한 지능형 교통안내 시스템 • 11
- 3 밀라노, 지능형 교통정보 시스템으로 신속·정확하고 손쉬운 길안내서비스 • 14
- 4 뉴욕주 시라큐스시, 데이터 분석을 기반으로 스마트 시티 추진 • 16
- 5 덴마크 베스타스 윈드 시스템, 풍력 에너지 관리로 에너지 생산 효과 극대화 • 18
- 6 구글, 실시간 자동 번역시스템을 통한 의사소통의 불편해소 • 22
- 7 월마트, 데이터 분석을 통한 투자수익 증대 • 24
- 8 자라, 정포별·상품별 등 실시간 데이터 분석을 통한 판매량 증대 • 27
- 9 마이크론 테크놀로지, 제품생산시간 분석을 통한 비용절감 • 30
- 10 코카콜라의 SNS 데이터 활용을 통한 가치향상 노력 • 33
- 11 리츠칼튼 호텔, 데이터 관리를 통한 고객맞춤형 서비스 제공 • 35
- 12 SNS를 활용한 할리우드 흥행 수익 예측 • 38
- 13 넷플릭스, 데이터 분석으로 온라인 DVD 판매제고 및 고객 서비스 향상 • 40

빅데이터로
진화하는 세상

Big Data 글로벌 선진 사례

53가지 사례



[국내편]

- 1 한국석유공사, 국내 유가 예보 서비스를 통한 비즈니스 최적화 • 42
- 2 국민권익위원회, 민원정보분석 시스템을 통한 국민과 정부의 소통 활성화 • 46
- 3 한국도로공사, 고객 목소리 분석 시스템을 통한 서비스 혁신 • 50
- 4 통계청, 임금근로일자리 통계로 일자리 현황 파악 지원 • 52
- 5 한국수자원공사, 스마트 워터 그리드를 기반으로 물 부족 현상 해결 • 54
- 6 포스코, 원료가격의 효율적 구매를 통한 가격 경쟁력 제고 • 57
- 7 GS EPS, 전력시장 분석 시스템을 통한 전력시장 전망 및 분석 • 60
- 8 SK텔레콤, 소셜네트워크에서의 여론분석을 위한 스마트 인사이트 시스템 • 62
- 9 현대·기아자동차, 문서 중앙화 시스템으로 통합적 기업 정보 관리 • 64

(빅데이터 커뮤니티, 자료실, http://www.bicdata.com/bbs/board.php?bo_table=business_archive&wr_id=88, 2012. 6)

빅데이터

빅데이터 커뮤니티
<http://www.bicdata.com>

빅데이터 국내외 적용 사례

Ⅱ 빅데이터를 통한 건강한 사회

[해외편]

- 1 미국 국립보건원, 유전자 데이터 공유를 통한 질병치료체계 마련 • 68
- 2 미국 국립보건원, Pillbox 프로젝트를 통한 의료개혁 • 72
- 3 미국 퇴역군인의 전자의료기록 분석을 통한 맞춤형 의료 서비스 지원 • 76
- 4 싱가포르, 주민위원회 센터 네트워크를 기반으로 맞춤형 복지사회 구현 • 79
- 5 캐나다 온타리오 공과대병원, 미숙아 모니터링을 통한 감염 예방 및 예측 • 81
- 6 건강보험회사 웰포인트, 슈퍼컴퓨터를 활용한 효율적인 환자치료 • 84
- 7 구글, 검색어 분석을 통한 독감예보 서비스 제공 • 88
- 8 네덜란드 스파크드, 빅데이터를 활용하여 건강한 소 사육 환경 구축 • 90

[국내편]

- 1 보건복지부, 행복e음을 통한 수요자중심의 복지서비스 구현 • 92
- 2 근로복지공단, '찾아가는 서비스'를 통한 맞춤형 고객관계 관리 • 95
- 3 한국인체자원은행네트워크, 정보공유로 생명공학분야 경쟁력 제고 • 98
- 4 분당 서울대병원, 임상 의사결정지원 시스템 등 의료 IT시장 선도 • 101
- 5 DNA Link, 유전자 분석 시스템으로 맞춤형 건강 검진 서비스 제공 • 103
- 6 연세대학교 의료원, 후(H∞H) 헬스케어 시스템으로 양질의 의료 서비스 제공 • 106

Ⅲ 빅데이터로 희망하는 안전한 사회

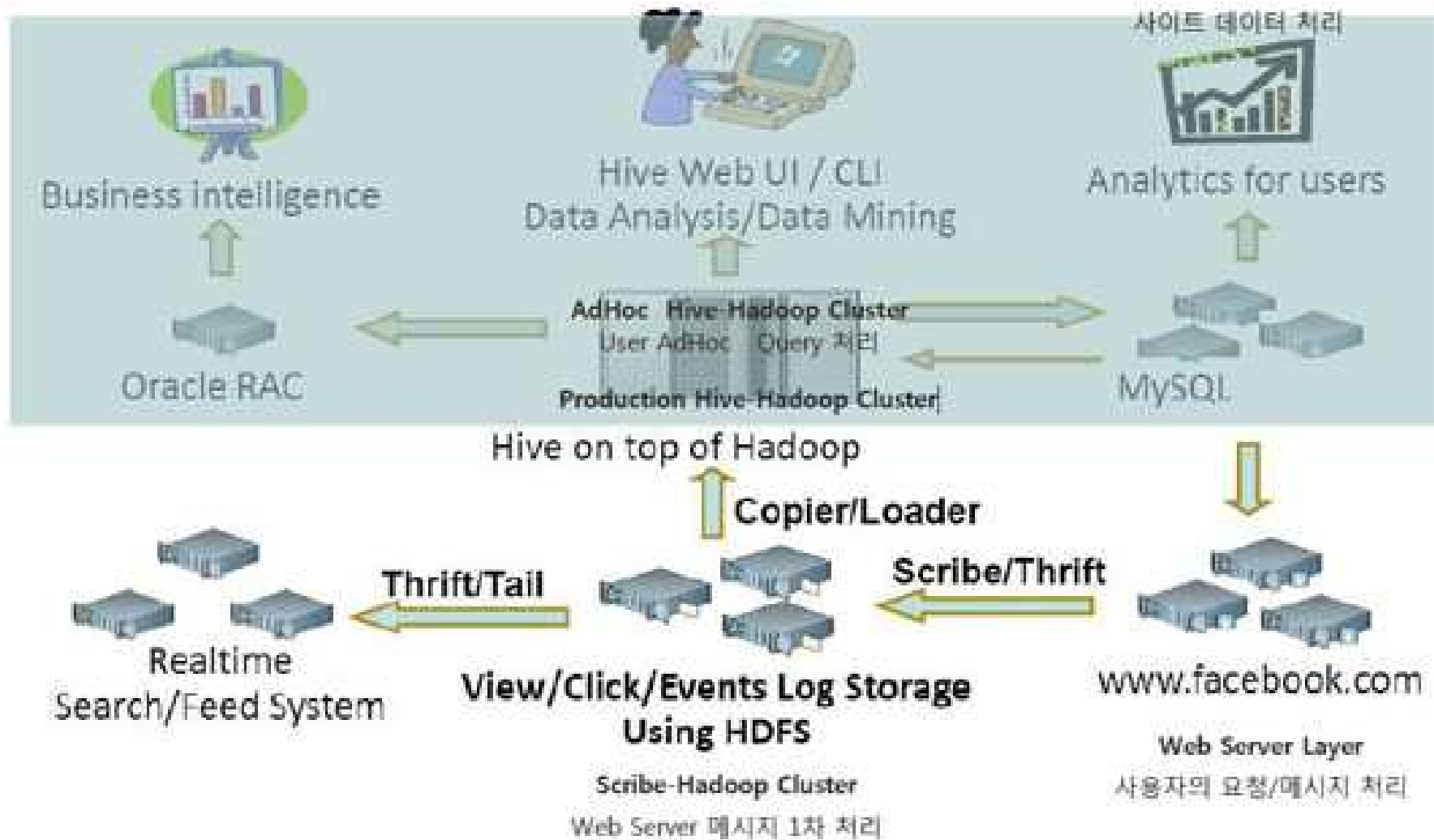
- 1 싱가포르, 국가위험관리시스템을 통한 국가안전관리 • 110
- 2 FBI, 유전자 색인 시스템 활용한 단시간 범인 검거 체계 마련 • 113
- 3 샌프란시스코, 범죄 예방 시스템으로 안전 지역사회 구축 • 116
- 4 싱가포르 출입국관리소, 통합적 정보분석으로 출입국 보안 및 국경 통제 강화 • 118
- 5 일본, 다양한 센서 데이터를 활용한 재난대응 능력 강화 • 121
- 6 위키리크스, 데이터 분석을 통한 효과적인 전술 제공 서비스 • 124
- 7 서울시, 시민과 함께 만드는 실시간 폭우지도로 수해 예방 및 대책 마련 • 127
- 8 KSTEC, 보험사기 방지시스템으로 보험사기 방지 • 129

Ⅳ 빅데이터로 소통하는 창의적 사회

- 1 "data.gov", 시민의 자발적인 참여와 창의성을 기반으로 정책수립에 기여 • 134
- 2 미국 미시간 주, 데이터웨어하우스 구축으로 공공서비스 질적 향상 • 138
- 3 영국 패치메이, 국민참여형 안전관리 플랫폼 구현 • 141
- 4 케냐 우사히디, 집단지성으로 이루어진 재난관리 오픈소스 플랫폼 • 144
- 5 IBM 왓슨, 인공지능 슈퍼컴퓨터로 인류의 창조성과 혁신 촉진 • 147
- 6 애플 시리, 지능형 음성인식을 통해 더 똑똑해지는 창의적 사고 가능 • 151
- 7 프랑스, 시민 건강 보호를 위한 스마트폰 소음지도 작성 • 154
- 8 새로운 과학적 발견: 대형강입자충돌기를 이용한 힉스입자 검출 • 156
- 9 솔트룩스, 부산자식네트워크 시스템으로 체계적 인적 네트워크 마련 • 158

(빅데이터 커뮤니티, 자료실, http://www.bicdata.com/bbs/board.php?bo_table=business_archive&wr_id=88, 2012. 6)

인프라 사례 - 페이스북 사용자 활동 분석

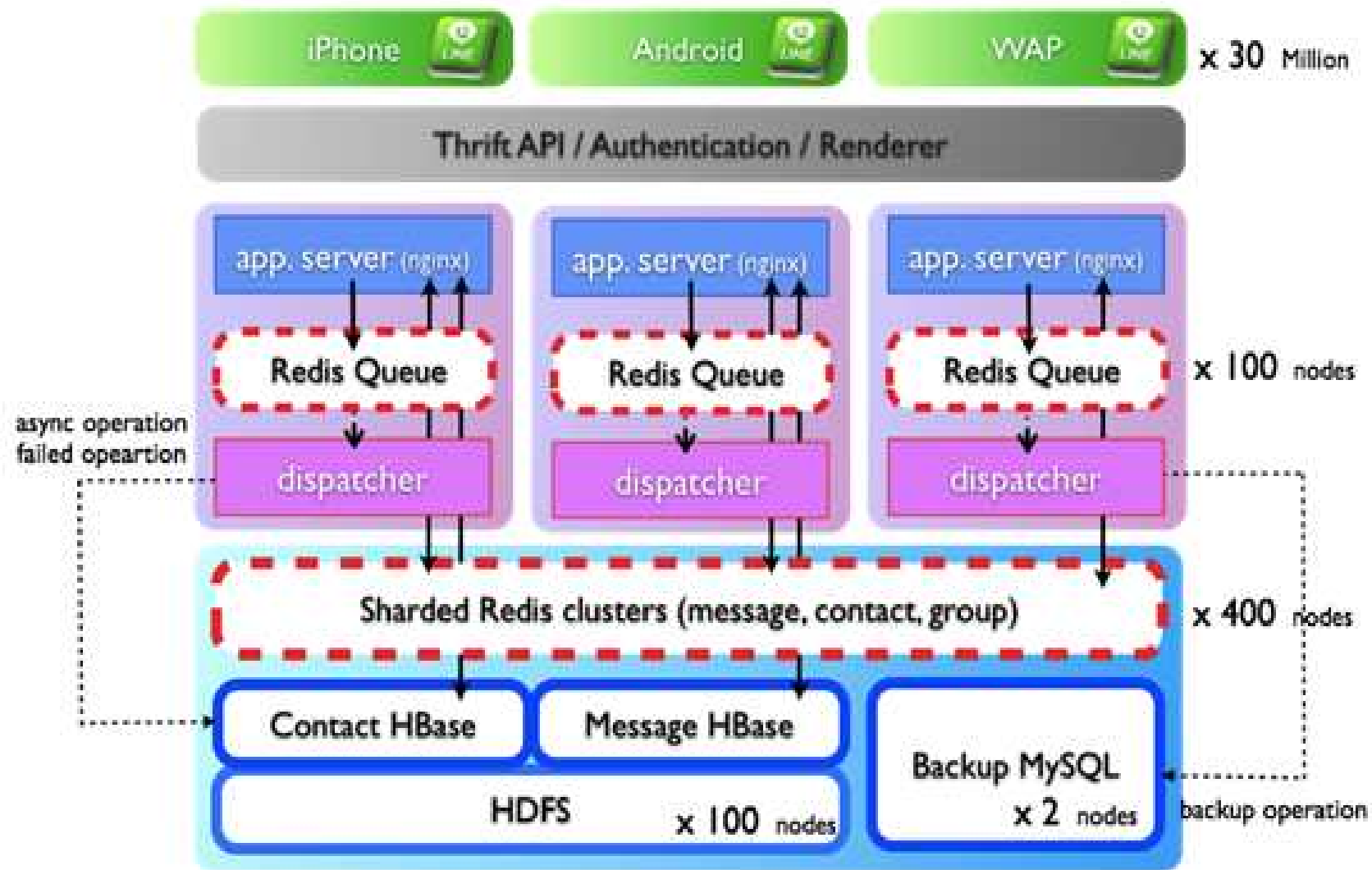


(투이컨설팅, 페이스북의 빅 데이터 처리 아키텍처)

빅데이터

빅데이터 커뮤니티
<http://www.bicdata.com>

인프라 사례 - NHN LINE Storage



(네이버재팬 엔지니어 블로그, <http://tech.naver.jp/blog/?p=1420>, 2012. 4)

빅데이터 커뮤니티 활동사항 - 학습



제1기 하둡개발자 그룹스터디



제2기 하둡개발자 그룹스터디



제2회 분석전문가 오픈세미나



1st 빅데이터 in 부산 세미나



제1회 빅데이터 플랫폼 기술세미나



제3기 분석전문가 그룹스터디

빅데이터 커뮤니티 활동사항 - 교육



충북대학교 빅데이터 대학원



한국RFID/USN융합협회



앵커웨이브 사내출장교육

“빅데이터 플랫폼 하둡 기초 강의” 커리큘럼

세부 교육내용

구분	시간	주요내용	강사
8. 6(월) (1일자)	09:00 ~ 11:00 (2시간)	○ 빅 데이터의 가치와 도입 전략 - 빅데이터 분석과 활용 가치 - 빅데이터 플랫폼과 오픈소스	이현종 대표 (Leesult)
	11:00 ~ 18:00 (6시간)	○ 빅데이터 플랫폼 구현 기술의 이해 - Hadoop의 이해 및 Hadoop 생태계 - HDFS(Hadoop Distributed File System)의 이해	장형석 대표 (닷컴솔루션)
8. 7(화) (2일자)	09:00 ~ 13:00 (4시간)	○ 빅데이터 병렬 처리의 핵심 기술 - MapReduce(대용량 분산/병렬 처리 프레임워크) - Hive(SQL 기반 Data Warehouse)	장형석 대표 (닷컴솔루션)
	14:00 ~ 18:00 (4시간)	○ 빅데이터의 처리와 분석 - Hadoop 설치 - Pig(데이터-호름 기반의 스크립트 프로그래밍 언어) - HBase(컬럼 구조의 저장소)	
8. 8(수) (3일자)	09:00 ~ 13:00 (4시간)	○ Hadoop 생태계의 핵심 4대 구성요소 - Flume(데이터를 수집해서 HDFS에 저장하는 프레임워크) - Sqoop(다양한 DBMS 자료를 HDFS로 입출력) - Oozie(Hadoop의 Workflow 스케줄러) - Zookeeper(분산 클러스터 시스템 관리 코디네이터)	장형석 대표 (닷컴솔루션)
	14:00 ~ 16:00 (2시간)	○ 빅데이터 적용 사례 - 국내외 Hadoop 적용 사례	이현종 대표 (Leesult)
	16:00 ~ 18:00 (2시간)	○ Hadoop 활용 솔루션 및 관리 시스템 - Chukwa(로그 수집 분석 및 HICC) - Hadoop Management	장형석 대표 (닷컴솔루션)

※ 사정에 따라 교육내용 및 강사는 일부 변경될 수 있습니다.



이현종 대표 / 빅데이터 커뮤니티

- 빅데이터 커뮤니티 운영자
- 검색엔진, 데이터 마이닝 전문
- 충북대학교 빅데이터 대학원 강사
- 빅데이터 자문 및 컨설팅



장형석 대표 / 닷컴솔루션

- 빅데이터 커뮤니티 운영자
- 검색엔진, 데이터 마이닝 전문
- 충북대학교 빅데이터 대학원 강사
- 빅데이터 자문 및 컨설팅



이동우 차장 / 지티원

- 빅데이터 커뮤니티 운영진
- 지티원 데이터거버넌스사업부 차장
- 빅데이터 분석플랫폼 with R 구축중
- 행안부 공공정보 신뢰성 방안 연구과제 수행

빅데이터 커뮤니티 마스터플랜

전문인력양성

빅데이터 하둡개발자 그룹스터디
빅데이터 분석전문가 그룹스터디



빅데이터 전문가그룹 워크숍
-> 파일럿 프로젝트 수행

교육프로그램

협회교육, 기업/대학/연구소 출장강의

개발자 이론과정, 실습과정
분석가 이론과정, 실습과정
기획자 이론과정



전문강사 확보
양질의 커리큘럼 개발
전문 인력 공급

세미나

빅데이터 오픈소스 플랫폼 기술세미나
빅데이터 분석전문가 오픈세미나



"빅데이터 산학연 포럼" 구성
-> 국가 정책에 의견 반영

테스트베드

기경험 활용한 테스트베드 구축
- 빅데이터 분석 플랫폼 제공



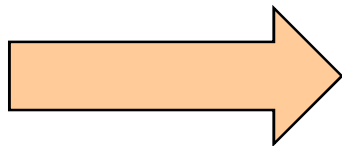
지식경제부 테스트베드 운영
-> 플랫폼과 분석가치의 조화

인프라구축

빅데이터 분석 인프라 구축 지원
- 정부기관, 연구소, 대학, 일반기업 등

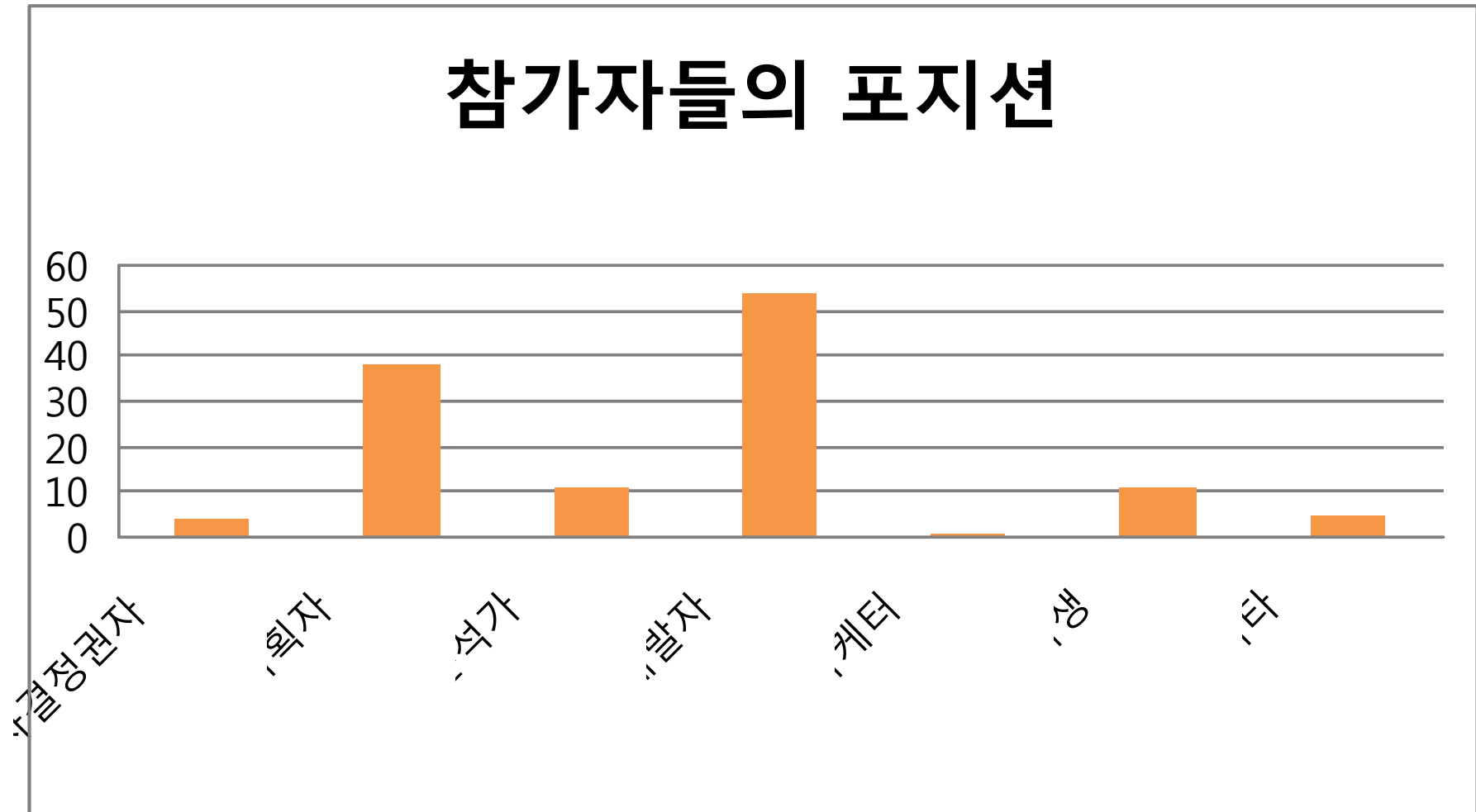


국가 빅데이터 산업의 조기 발전
-> 전문인력의 선순환 활용



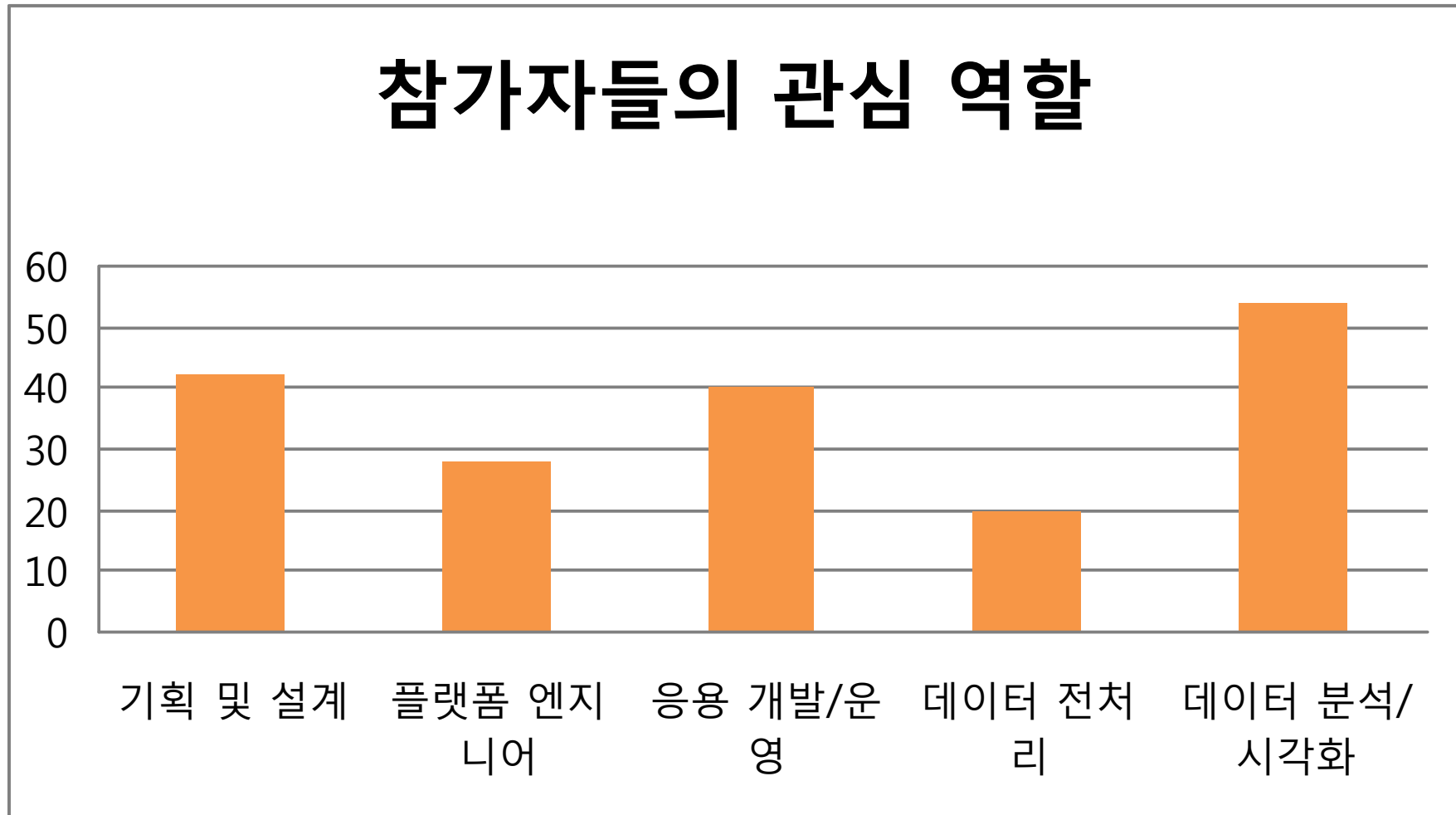
정보 공유를 통한 국내 빅데이터 발전에 기여

첨부 - 제1회 기술세미나 설문결과



(제1회 빅데이터 오픈소스 플랫폼 기술세미나, 124명 중복 답변, 2012. 7)

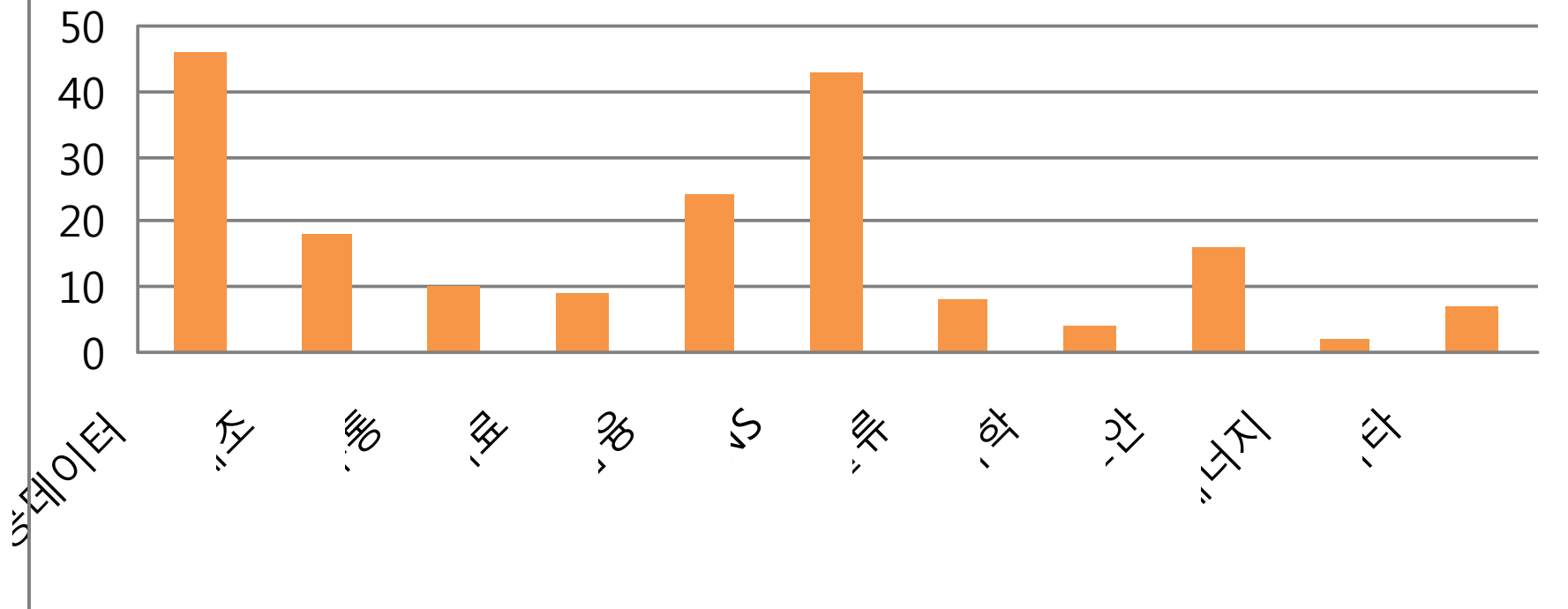
첨부 - 제1회 기술세미나 설문결과



(제1회 빅데이터 오픈소스 플랫폼 기술세미나, 124명 중복 답변, 2012. 7)

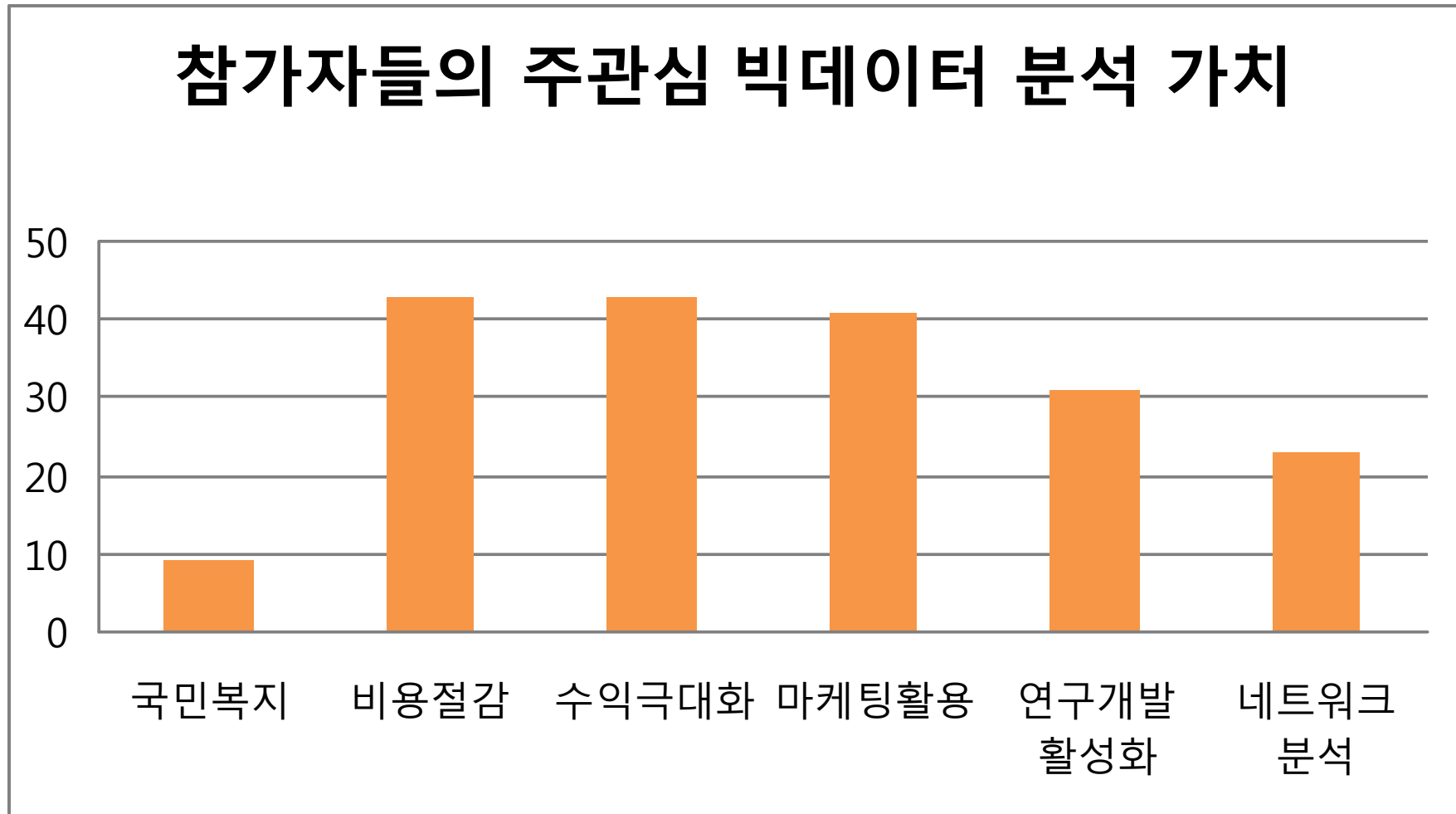
첨부 - 제1회 기술세미나 설문결과

참가자들의 주 관심 데이터 분야



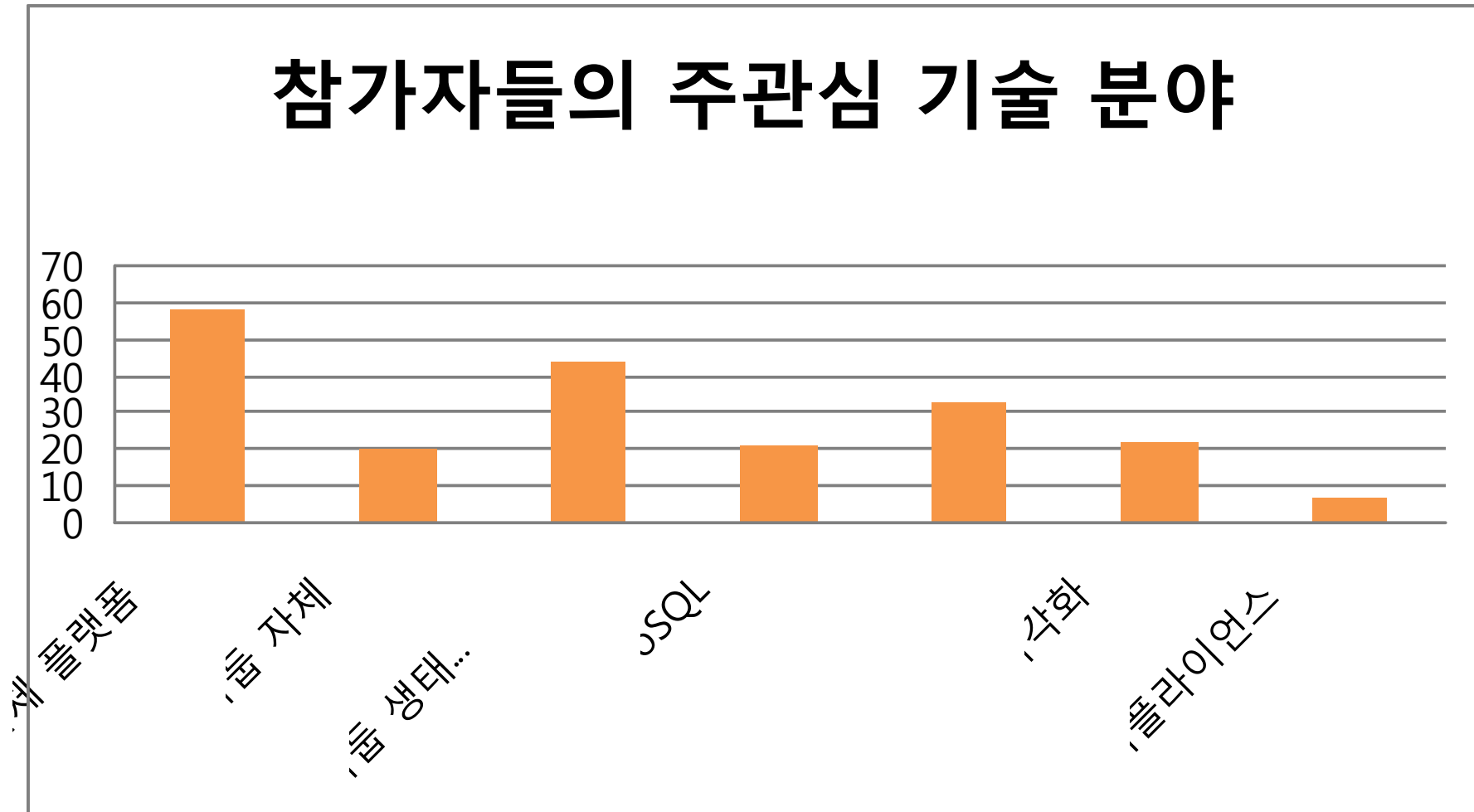
(제1회 빅데이터 오픈소스 플랫폼 기술세미나, 124명 중복 답변, 2012. 7)

첨부 - 제1회 기술세미나 설문결과



(제1회 빅데이터 오픈소스 플랫폼 기술세미나, 124명 중복 답변, 2012. 7)

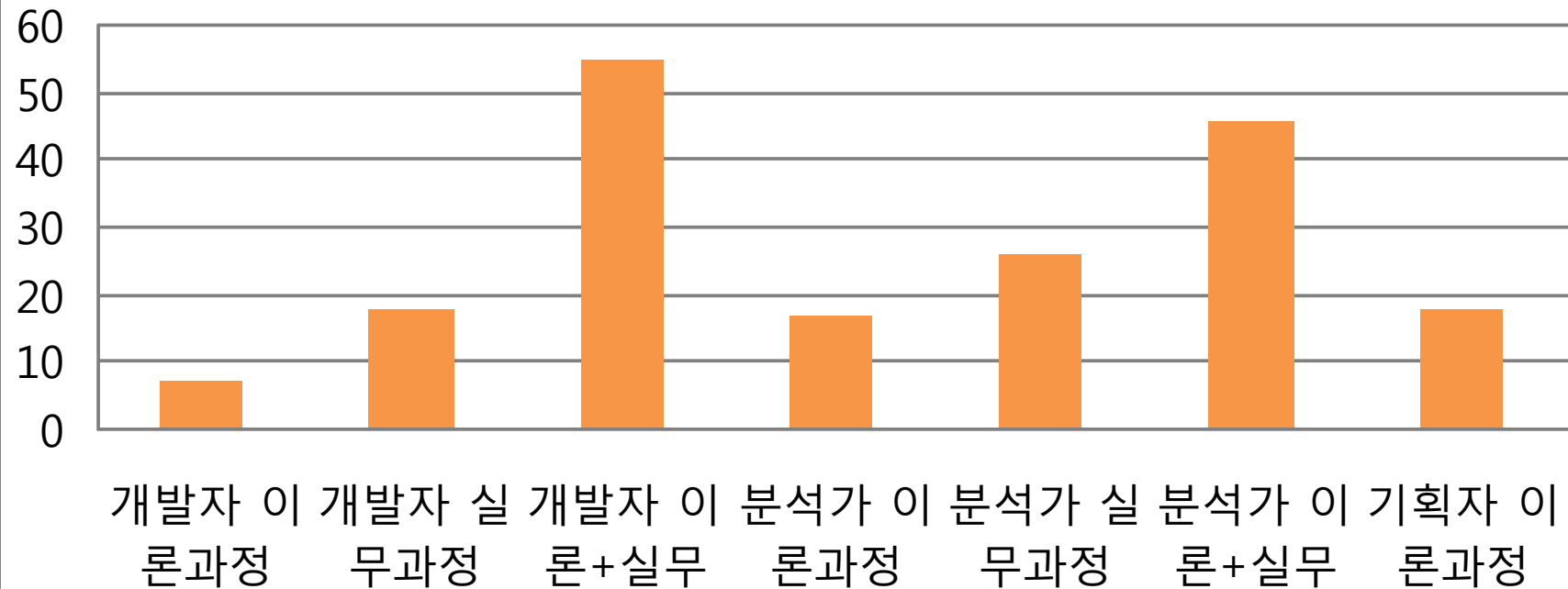
첨부 - 제1회 기술세미나 설문결과



(제1회 빅데이터 오픈소스 플랫폼 기술세미나, 124명 중복 답변, 2012. 7)

첨부 - 제1회 기술세미나 설문결과

참가자들의 희망 교육과정



(제1회 빅데이터 오픈소스 플랫폼 기술세미나, 124명 중복 답변, 2012. 7)

감사합니다



leesult@gmail.com