

빅데이터와 수자원정보화

중부대학교 토목공학과
이 주 헌

CONTENTS

1. 빅데이터의 정의
2. 빅데이터 사례
3. 국가수자원 정보화
4. Kwater 빅데이터기반 물정보화
5. 우리의 과제

1. 빅데이터의 정의

향후 쓰레기에서 근사한 관광자원을 발굴하듯이
폭발적으로 증가하는 데이터에서 새로운 가치 창출하는 것이 과제임



와인 빈티지와 빅데이터



Copyright(c)2013 조성준

- 와인품질은 좋은 저장고에서 대략 10개월 숙성시킨 후 전문가 품평회에서 산정
- 미국 프린스턴대 경제학과 아센펠터 교수는 프랑스 보르도 지방 **날씨를 이용하여 와인품질공식 개발** : 술을 담근지 3개월도 안돼 와인가치 예측

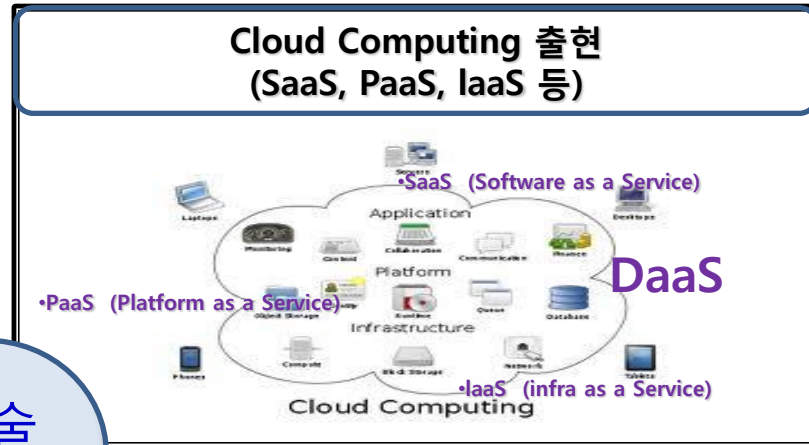
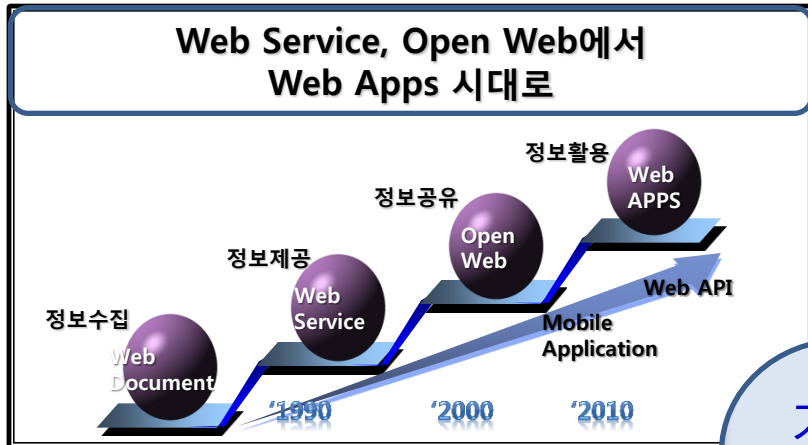
- 와인 품질 = $12.145 + 0.00117 \times \text{전년도 겨울 강수량} + 0.06140 \times \text{당해년도 평균기온} - 0.00386 \times \text{수확기 강수량}$
 - 예상과 달리 강수량보다 평균기온이 500배 더 영향
- 평론가들과 달리 1986년산 보르도 와인보다 1989, 1990년산 와인이 더 품질이 좋다고 예측. 결국 아센펠터의 예측이 맞았던 것으로 밝혀짐.

[통계] 가설 → 데이터수집 → 검증

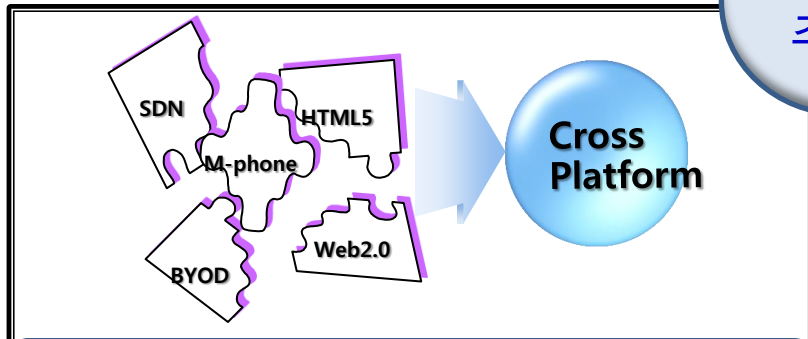
[빅데이터] 데이터분석 → 지식발견 → 검증 : **원인과 결과가 아닌 상관성**



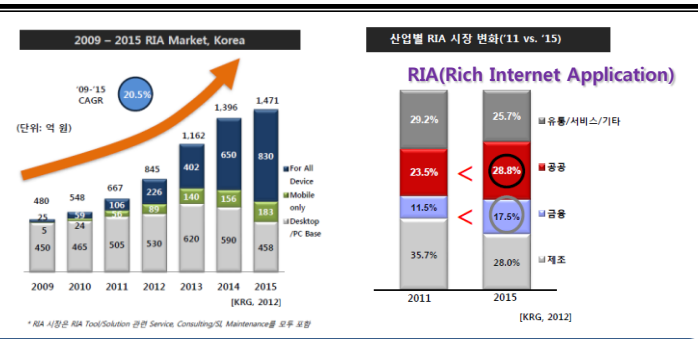
Web Apps(정보활용), Cloud Computing(> DaaS로 진화), Cross Platform(웹 브라우저), UI/UX(Visualization) 강조



기술 추세



Web Brower 기반의 Cross 플랫폼시대로



Mobile 사용확대 및 Web Site UI/UX 수요증가에 따른 RIA 시장 확대

빅데이터를 통해 미래를 알 수 있기 때문입니다.

- 공학기반의 공식이나 데이터 해석을 수행하는 이유는 이를 통해 미래를 알 수 있기 때문입니다.
- 빅데이터를 사용하는 이유도 미래를 알고 싶기 때문입니다.

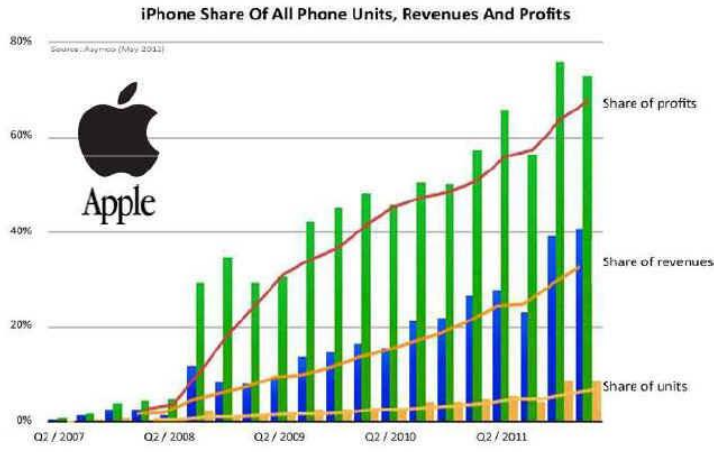
최근 너무 많은 것들이 달라지고 있기 때문입니다.

- 지구온난화에 의한 기후변화로 과거 50년동안 축적된 지식이 점점 의미를 잃어가고 있습니다. (연평균 강우량에 기초한 물수지분석, 과거 최대홍수량에 기초한 제방고 등등)
- 우리 삶에 밀접한 너무나 많은 것들이 빠르게 변하고 있습니다. 전문가로써 우리는 무엇인가를 말해야 하는데 옛 지혜와 스승은 여기에 없습니다. - 새로운 지도를 그려야 할 때가 온 것 같습니다.

경제적으로 연구를 수행하고 효과적으로 성과를 낼 수 있기 때문입니다.

- 구글, 아마존, 바이두 등 최근 많은 부를 축적한 기업이 빅데이터를 기반으로 하고 있습니다.
- 왜일까요? 경제적이고 합리적이기 때문입니다.

- 생산의 10%, 매출의 40%, 이익의 70%



Apple의 성공요인 ?

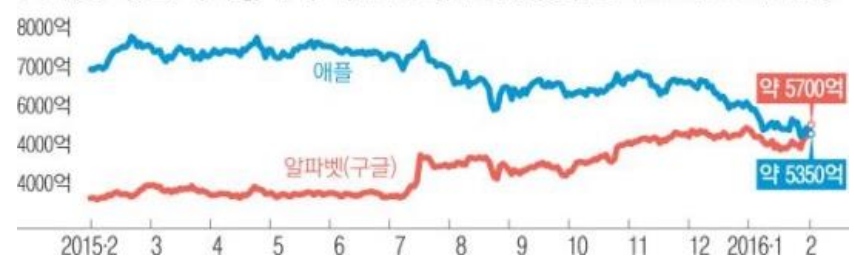
- HW 제조회사로서 SW 능력 보유
- Samsung이 Android를 인수했다면 ?

2012년

비교항목	애플	구글	MS(마이크로소프트)
주가	244.11 달러	474.47 달러	25.01 달러
주식발행수	9억0994만주(909.94M)	3억1849만주(318.49M)	87억6369만주(8.76B)
시가총액	2220억7000만달러 (약 270조원)	1514억 3244만 달러 (185조원)	2191억8000만달러 (267조원)

삼성전자: 909억 달러

알파벳·애플 시가총액 추이 단위:달러, 자료:삼성증권, 블룸버그 ※시간 외 거래 반영

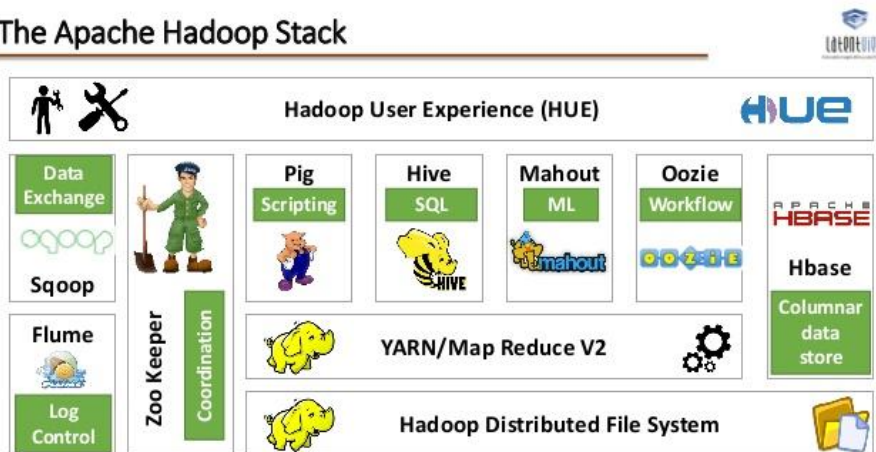


삼성전자: 2200억 달러

빅데이터 기술은 “Like Google” 입니다.

- 과거 오라클(RDBMS)의 시대에서는 접근하기 어려웠던 문제를 Google 방식(GFS)으로는 가능해졌습니다.
- Like Google 이란 수천대의 컴퓨터를 동시에 사용한다거나, 수천 테라바이트의 데이터 속에서 의미 있는 정보를 찾는다거나, 연관성 있어 보이는 것들은 서로 연결하는 기술과 방법입니다. (알파고-인공지능)
- 빅데이터는 구글에서 시작했다고 봐도 좋을 것 같습니다. 그렇다고 구글 것을 마음대로 사용할 수 없겠지요. 그래서 수많은 어플리케이션이 만들어졌습니다. - **Hadoop Stack, ELK Stack, Splunk Stack**

The Apache Hadoop Stack



과거에는 왜 빅데이터를 몰랐을까요?

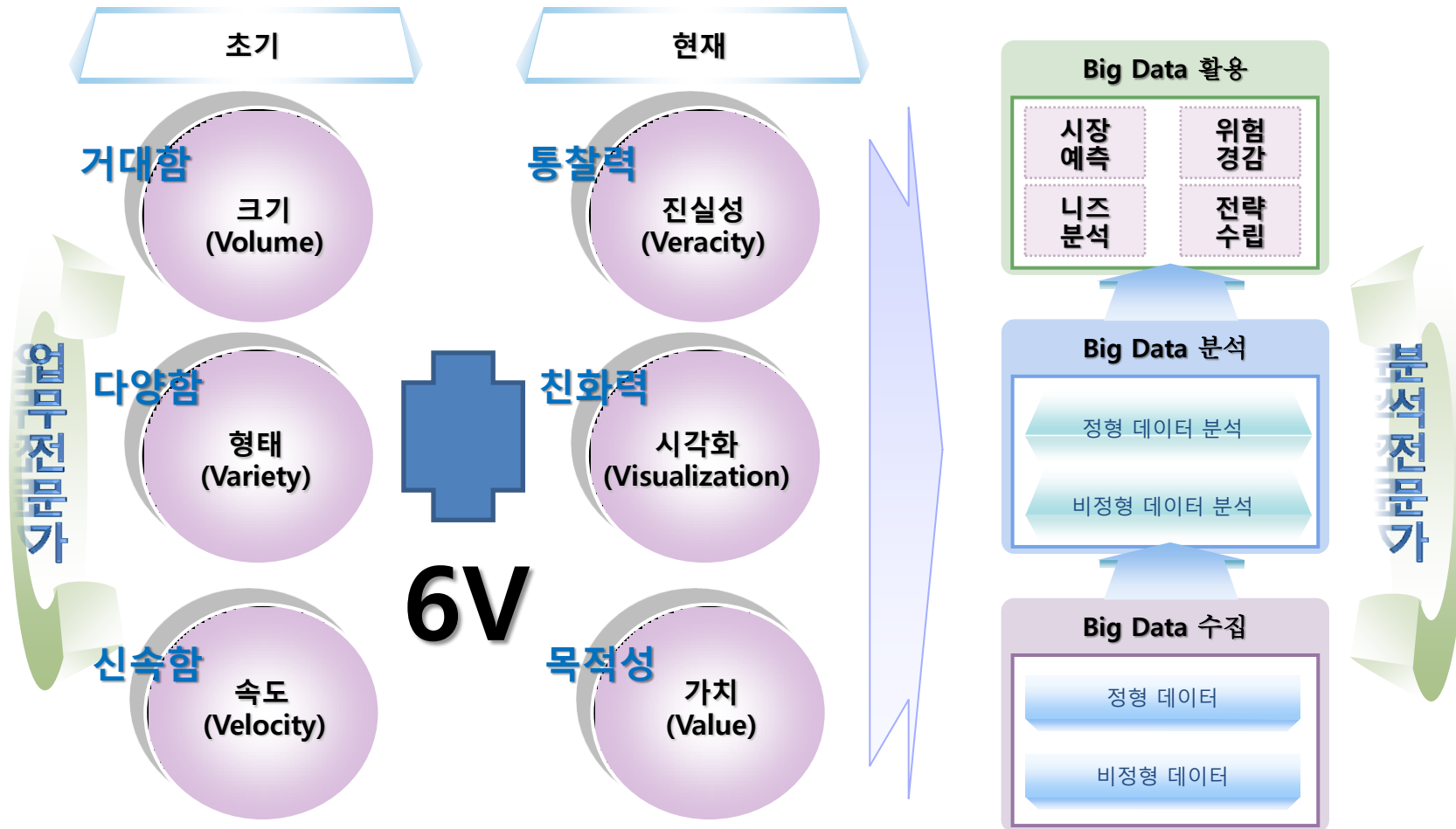
1. 빅데이터의 정의

첫째, 지금까지는 사람이 어떻게 행동했다, 물류가 어떻게 이동했다, 재고량이 어떻게 증감했다 등의 모든 사건에 대한 **저장기술이 기술적으로 불가능**했다는 것이다. 하지만 디지털 기술의 발달과 디지털 장치의 확산으로 이 모든 사건을 기록할 수 있게 되었다.

둘째, 지금까지는 디지털 빅데이터를 저장하는 것이 **경제적으로 합리적이지 않았기** 때문이다. 1980년대에는 1Gbyte를 저장하기 위해서 100만\$(10억원 이상)가 드는 등 경제적 부담이 컸다. 하지만 2010년도에 들어서면서 0.1\$(100원) 미만으로 떨어져 가격적으로 그리 부담스럽지 않게 되었다.

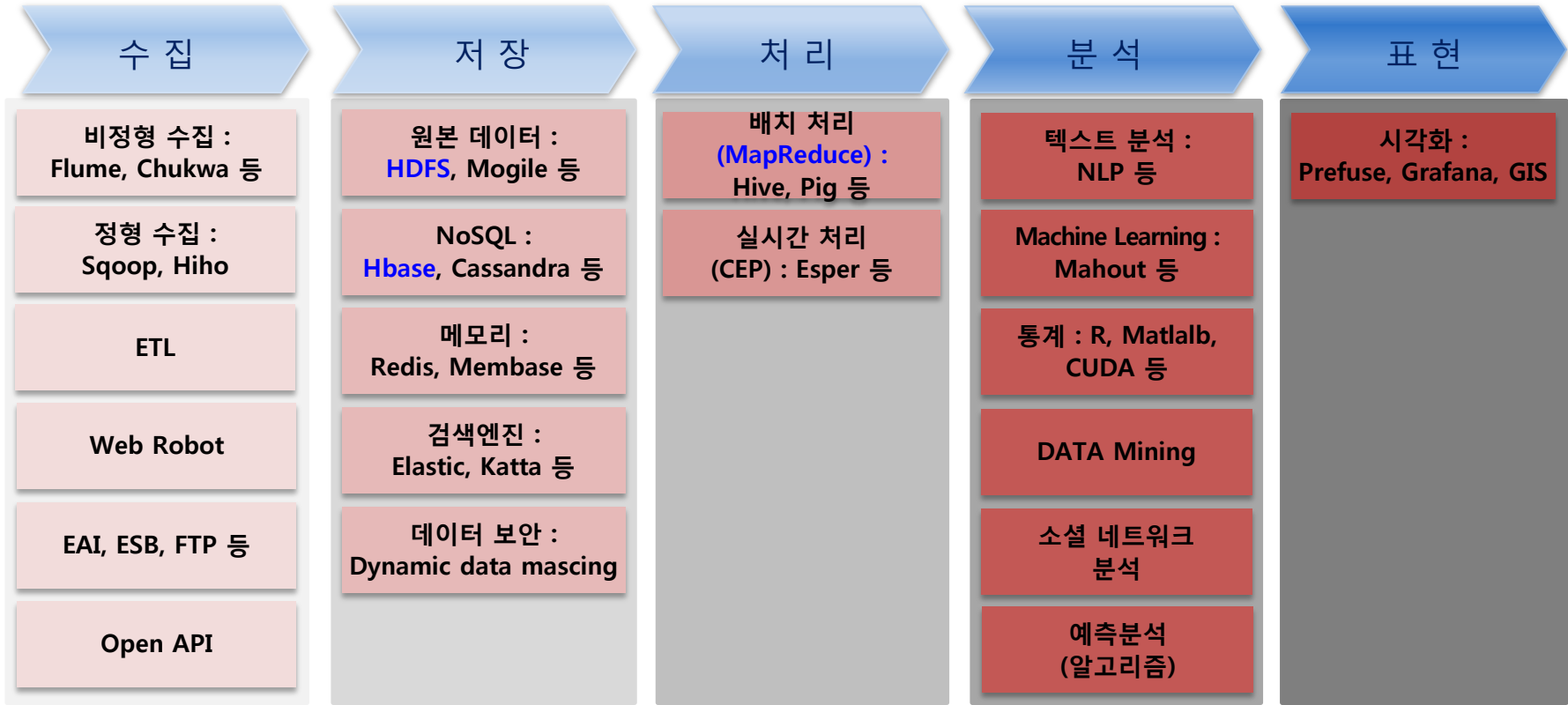
셋째, 과거에는 하루에 발생하는 엄청난 양의 데이터들을 한 대의 컴퓨터로 읽는 데만 해도 엄청난 시간이 소요된다는 문제가 있었다. 하지만 지금은 이러한 과정이 **병렬 분산처리기술(클러스터)** 등의 기법이 활성화 되면서 가능해졌다.

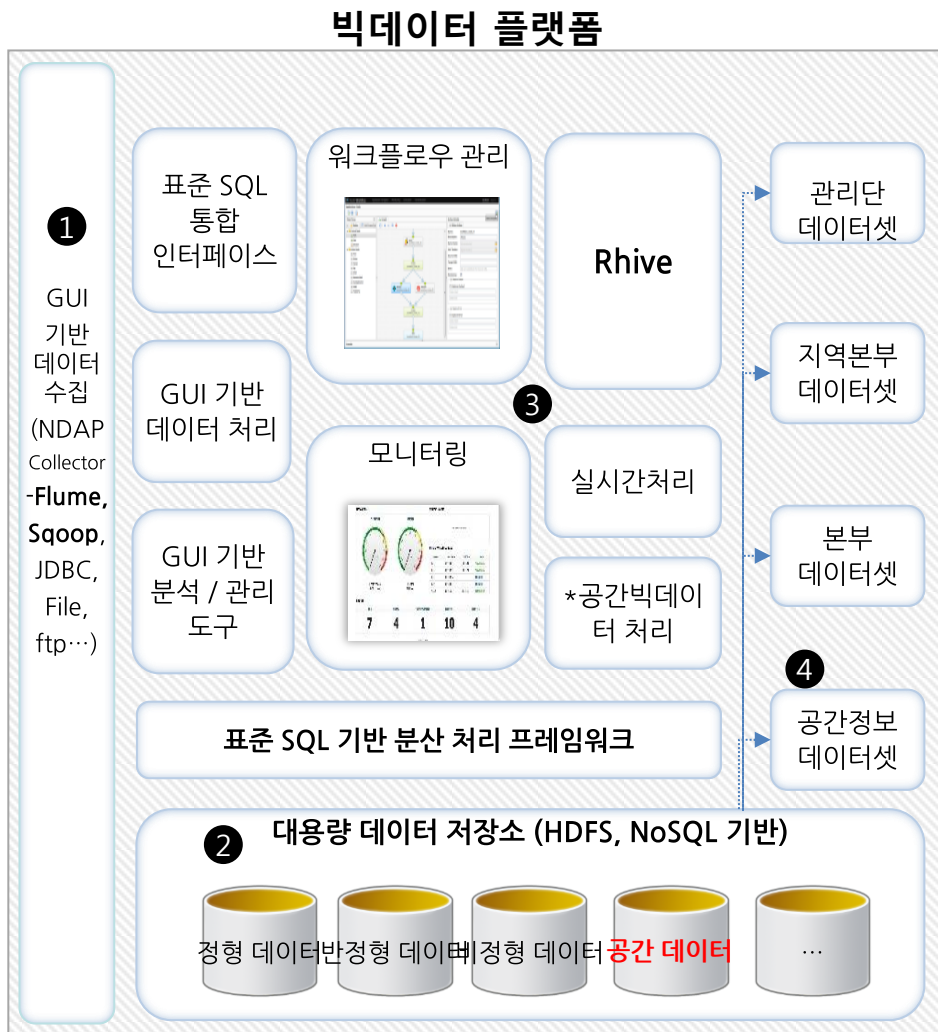




데이터 수집, 저장, 처리, 분석 및 표현에서 많은 솔루션들이 출현하였고
 해외에서 분석 및 시각화 플랫폼으로 Splunk 및 ELK Stack가 각광

● 대용량 데이터 기술





- ① 데이터의 수집
- ② spatial-framework-for-Hadoop 설치 / 공간 Feature (Point, Line, Polygon) 포함한 공간 데이터를 **HDFS**에 저장
- ③ **Hive**에서 SQL 쿼리로 공간데이터 조회 / 실시간 자료처리
- ④ 공간정보 데이터 셋 구축
- ⑤ R package 혹은 ArcGIS에서 공간정보 시각화

*Esri GIS tools for hadoop : Hadoop 기반에서 빅 데이터 공간정보 애플리케이션 작성을 위한 오픈소스 프레임워크

2. 빅데이터 사례

현재 해외 글로벌 IT기업들은 빅데이터를 차세대 주력분야로 선정

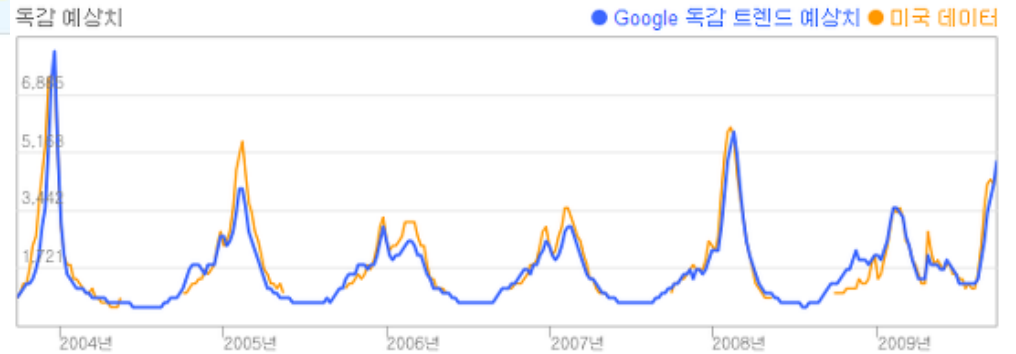
해외 주요기업 빅데이터 사업 추진현황			기타 해외기업 사업 추진현황
구글	MS	EMC	<ul style="list-style-type: none"> Teradata: 비정형 데이터의 고급분석 관리 솔루션 업체인수(애스터데이터) SAS: IT+분석+비즈니스 통합 플랫폼 구현 Oracle: 세계적인 DB업체 '하이페리온' 인수 IBM: 비즈니스 분석업체 인수 (5년 140억 투자) HP: 솔루션, 기업용 검색엔진 업체 인수
<ol style="list-style-type: none"> 대용량 처리기술 발표 - GFS, MapReduce 등 빅쿼리 서비스 공개 - 이용자 업로드 데이터 분석 서비스 	<ol style="list-style-type: none"> 윈도 애저 & 서버 플랫폼용 아파치 Hadoop 개발 계획 Hadoop 기술 전문 업체 호튼웍스와 협력 	<ol style="list-style-type: none"> 그린플럼, 아이실론 등 데이터 관련 업체인수 - 데이터 저장, 관리, 분석 제공 빅데이터 스토리지 솔루션 제공 - 아이실론, 아트모스, 다큐멘텀 등 EMC 애널리틱스 랩 운영 - 데이터 사이언티스트 육성 	

해외 글로벌 기업들은 빅데이터 전문기업의 인수합병 등을 통해 선도적 전략을 수립함과 동시에 빅데이터를 접목시켜 고급분석을 강화하는 전략 위주

Google does faster than CDC (<http://www.cdc.gov>)



미국 독감 유행 수준
독감 예상치



미국: [미국 질병 통제 센터](#)에서 공개한 인플루엔자 의사환자(ILI) 데이터

● NOAA / NWS

- 매년 30 페타바이트의 신규 데이터 관리
- 매일 35억개의 위성, 선박, 항공기, 부표 등의 센서에서 데이터 수집
- 대기, 해양, 지상 관측 데이터를 수집하여 정확도가 높은 예측 모델링

Time-Series Analysis

● Random Signals and Noise

Look at distributions, autocorrelations and other statistics of various types time-series. Time-series types include gaussian distributed, exponential, and white. Interactive javascripts all observe the effects of changing different parameters. From the Georgia Institute of Technology School of Electrical and Computer Engineering course in random signals and noise.

● Wavelet Analysis

Many time series are "non-stationary," varying in both amplitude and frequency over long periods of time. Wavelet analysis is a technique that decomposes a time series into time/freq simultaneously. By doing so, one can get information on both the amplitude of any "periodic" signals within the series, and how this amplitude varies with time. This interactive page all plot these amplitudes of a selected time series at various periodicities through time. Several ocean and atmospheric time series are available.

Climate Tools

● UV-CDAT Climate Data Analysis Tools

U.S. Department of Commerce | National Oceanic & Atmospheric Administration

NOAA Earth System Research Laboratory
Physical Sciences Division

Physical Sciences Division About Contact Research

Climate Datasets: By Category

- All
- Sub-daily
- Daily
- Monthly
- Surface
- Temperature
- SST
- Precipitation
- Land
- Ocean
- Multi-level
- Radiation
- Arctic
- Reanalysis
- Climate Indices
- Search Datasets 🔍
- 20th Century Reanalysis
- Popular Datasets**
- ICOADS
- NCEP/NCAR Reanalysis
- N. American Regional Reanalysis
- Plotting & Analysis**
- Basic Plots
- Analysis Tools

A sample plot is shown on the right o

Station Data

- **US and Global Station Data**
- Use **Climvis** to plot time-series of station temperature (maximum, minimum and precipitation and Palmer Drought Inde
- **Western Regional Climate Center**
- Western Regional climate center station rainfall probabilities and related quanti

Climate Datasets

- **Climate Reanalyzer**
- Climate Reanalyzer is being developed. Investigate climate using interfaces for map layers to Google Earth.

기상기후 빅데이터(Big Data) 기상청 Korea Meteorological Administration

빅데이터

기존 방식으로 **저장/관리/분석하기 어려울 정도의 큰 규모**이며 유통 속도가 빠르고 형태가 다양해서 복잡하지만 큰 가치를 갖는 데이터 집합


기상기후 빅데이터

방대한 기상기후 관련 **데이터**, 과학적 데이터 **분석 기법**, 다양한 분야와 접목한 기상기후 융합 **서비스**를 통칭

데이터 저장

10PB 이상

수치모델, 관측(위성, 레이더), 기후통계, 해외기상 등




데이터 분석

- 생활·산업·보건 기상 지수
- 감기/자외선/식중독 등
- 계절기상 서비스
- 봄꽃/단풍/김장 등
- 지역기후서비스
- 제주감귤, 전남 갯벌지수 등
- 응용기상서비스
- 수문, 스포츠, 항공, 농림, 도로 등
- 기후자료서비스
- 기후통계사, 리오, 세계기후지수 등

데이터 개방

[2013년 기준]



중앙부처 개방률	기상DB 개방률	기상DB Open API 사용자
21.6%	80.6% (25/31)	24% (6/25)

* 1PB(페타바이트): 일반적인 노래 파일 기준 2억곡 분량에 해당

자료출처: 기상청 “기상기후 빅데이터포럼 활동보고서” 빅데이터 활용 기상융합서비스

기상기후 빅데이터 기반 추진과제

- 기상기후 빅데이터 분석 플랫폼 구축 및 외부 개방(15년~)
 - 대용량 기상기후 자료를 직접 분석할 수 있는 기술 기반 환경 마련 및 단계적 외부 개방
- 공공데이터포털, 타 기관 빅데이터 플랫폼 간 데이터 공유 기반 마련을 통해 데이터 셋 확보(15년~)
 - 기상기후-타 분야 데이터 기반 빅데이터 분석 경진 대회 공동 개최 등을 통한 활용 확산



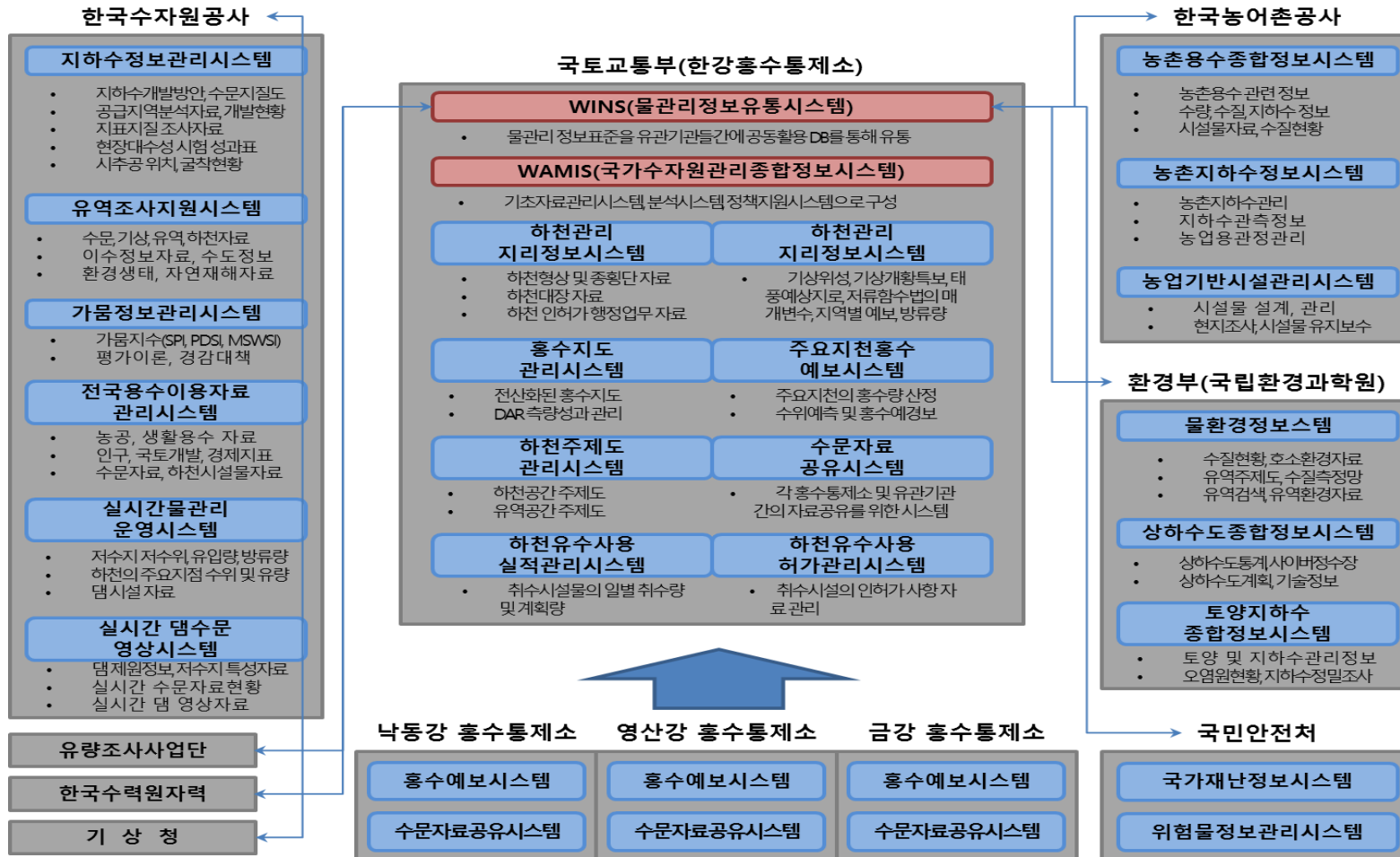
중점과제 및 차 순위과제 추진 계획

과제명(16개)	분과(9개)	추진유형/예산
별 상세 기상정보 서비스	체육관광 교통물류	정보화 / 약 7억원
2014 기상정보 융합 해수욕장 관광수요 예측 서비스 개발	농수축산	
2014 기상기후 민감 직물의 주산지 기상정보 서비스 및 생산성 분석	환경	연구개발 / 약 25억원
2014 기상정보 융합 한국연안 적조의 원인종과 밀도 변동 예측 연구	에너지	
2014 기상·재해에 따른 전력설비 피해 예측 및 조기경보 시스템 구축	수자원	
2014 기상기후 정보를 활용한 민간 에너지 효율화 및 수요 예측 서비스	체육관광	
2014 빅데이터 기반 현장기동지수 개발	보건	
2014 수도권 상세 홍수·침수 위험지역 예측 플랫폼 구축	교통물류	
2014 평창동계올림픽의 선수와 지도자를 위한 기상기후 정보앱 서비스	방재	
2014 기상기후정보를 활용한 맞춤형 생활체육지수 개발	산업	
2015 기상기후 민감형 감염병 발생 감시체계 및 발생 예측 시스템 개발		
2015 고혈압 환자의 응급 합병증 발병 기상위험 예측		
2015 지자체 기상상태별 교통사고 위험도 예측 서비스		
2015 도로기상 데이터 수집 및 활용체계 구축		
2015 차량용 강우센서를 이용한 도시에서의 실시간 도로통제 시스템 개발		
2015 빅데이터를 활용한 맞춤형 산업기상지수 개발		

자료출처: 기상청 "기상기후 빅데이터포럼 활동보고서" 빅데이터 활용 기상융합서비스

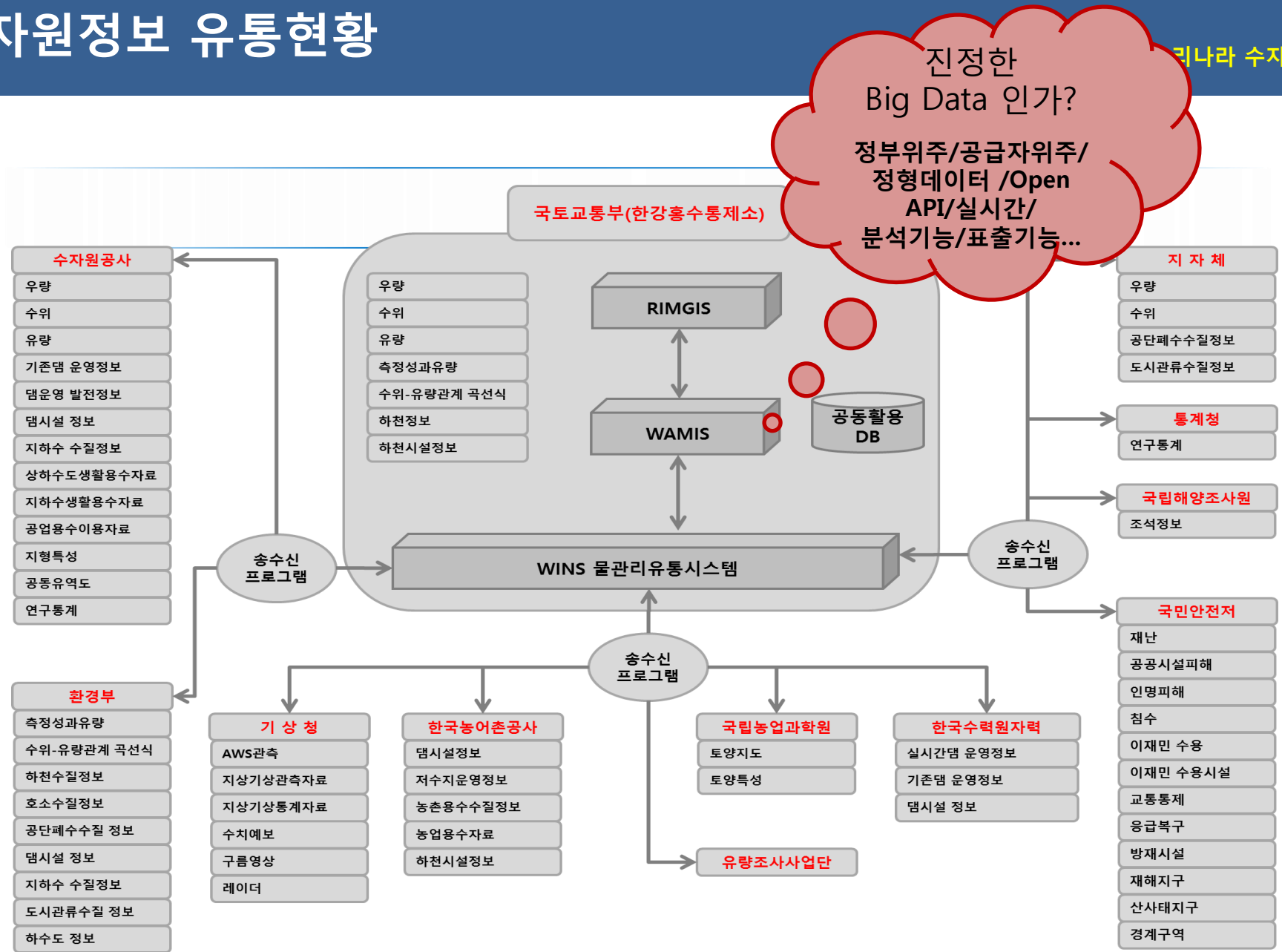
3. 우리나라 수자원분야 정보화

국토교통부 산하 한강홍수통제소를 중심으로 관계부처와의 시스템 연계 및 정보 제공을 통하여 관리 되고 있으며, 수질은 환경부, 재해관리는 국민안전처가 관리



수자원정보 유통현황

우리나라 수자원분야 정보화



My Water
K-water와 함께하는 물정보포털

검색어를 입력하세요



물과생활	물과지식	세계물정보	정보마당	재해안전	물과통계	운영권측	아플리케이션	전체보기
물과어플	물백과사전	글로벌자료원	이슈리포트	기후변화	쉬운물통계	수자원	Web App	
물과문화	물교육	제7차세계물포럼	최신기술동향	홍수	우리나라물통계	상수도	Eng App	
생활속의 물	지식동영상	여행자 먹는 물	학습정보	가뭄	세계물통계	하수도		
갤러리	물과역사	세계물유기	물관리기관/단체	수질/오류		지하수		
나의물이야기		K-water 해외사업	소셜&뉴스	기상정보				
			수병&정백					
			물행사양정					
			Q&A					

WAMIS
국가수자원관리종합정보시스템

기초자료검색 : 물지로 : 모두미 : English : Sitemap

수문기상 | 유역 | 하천 | 댐 | 지하수 | 이수 | 수도 | 환경생태 | 자연재해 | 지형공간

뉴스

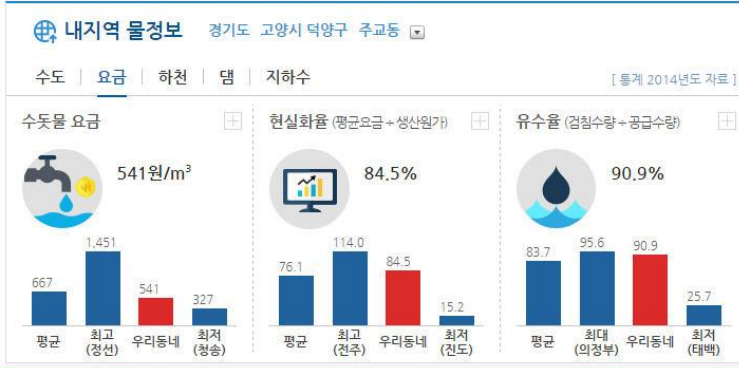
제주, 지하수 수질 1등급 지역 환경개발 불허 | 2016.07.14 | 뉴스시스

625억 돌인 4대강 첫 도수로 한달 쓰고 운... | 2016.07.14 | 연합뉴스

오늘도 뚝뚝더위... 동해안 비, 중부·경상내륙... | 2016.07.14 | 뉴스시스

김해시 수돗물 불소 투입 '오락가락'... 행정... | 2016.07.13 | 연합뉴스

장충군 평촌천 물난리 걱정 '뚝'... 국비 54억... | 2016.07.13 | 뉴스시스



행사/일정

국제행사 0

2016 Singapore

기타 0

대한전기학회 추

기타 0

2016 한국약학:

나의 물이야

물 여를 유가...

공간도형 다운로드

지도로검색 바로가기

USE GUIDE
WAMIS 정보이용 길잡이

MOBILE APP
SMART 물정보 앱서비스

자료실

통계도 보는 한국의 수자원 2014-04-03

겨간지 '수자원' 불호 2012-06-04

겨간지 '수자원' 겨울호 2012-03-06

물관리 정보의 모든것

Water Resources Management Information System
국가수자원관리종합정보시스템

물관리 정보를 대국민 서비스하기 위해 구축 운영되고 있는 인터넷기반의 포털시스템입니다.

새소식 | 게시판

[공지] 2016년 WAMIS 홈페이지... 2016-01-18

[공지] 2016년 WAMIS Smart물경... 2016-01-18

[공지] 용수업리미 업 실행하고... 2015-06-16

[공지] 2015년 WAMIS 서비스 기... 2015-06-16

[공지] 기린정보보호법에 따른... 2012-03-30

[공지] WAMIS 개선사항 알림 2012-02-09

강수량자료 | 수위자료 | 기상자료

댐수문자료 | OPEN API | 지도로 검색

검색영역선택

구분 유역별

대련여 =선택하세요=

중련여 =선택하세요=

지도로검색 2016년 7월 13일


물지도보기

토막 수자원 정보

백수로

천수의 침입을 하고, 지구나 육은 언젠 배후지의 화

영수를 저조 시켜 바다로 배격하기 위한 시설



Earth System Research Laboratory

Physical Sciences Division

Search PSD:

[Calendar](#) | [People](#) | [Publications](#)

Physical Sciences Division

[About](#) [Contact](#) [Research](#) [Data](#) [Products](#) [News](#) [Outreach](#)

Climate Datasets: By Category

- All
- Sub-daily
- Daily
- Monthly
- Surface
- Temperature
- SST
- Precipitation
- Land
- Ocean
- Multi-level
- Radiation
- Arctic
- Reanalysis
- Climate Indices
- Search Datasets
- 20th Century Reanalysis
- Popular Datasets**
- ICOADS
- NCEP/NCAR Reanalysis
- N. American Regional Reanalysis
- Plotting & Analysis**
- Basic Plots
- Analysis Tools
- Access**
- FTP Access
- OpenDAP Access
- Software Resources**
- "Complete Data Resources"
- About NetCDF
- Reading Data in NetCDF Format
- FAQ**
- Other Data Resources
- Acronyms
- How to Cite
- Feedback

Other Climate Analysis and Plotting Webtools

Station Data:	US and Global <input type="checkbox"/> Western States <input type="checkbox"/> Southeastern States <input type="checkbox"/>
Climate Datasets:	Climate Reanalyzer TAO Buoy Data <input type="checkbox"/> PMEL's LAS Server <input type="checkbox"/> IRI/LDEO <input type="checkbox"/> EPIC <input type="checkbox"/>
Time-series Analysis:	Random Signals and Noise <input type="checkbox"/> Wavelet Analysis <input type="checkbox"/>
Climate Tools:	NEW CDO CDAT Climate Tools <input type="checkbox"/>

A sample plot is shown on the right of each package description. Click on the plot for the full size version.



Station Data

- **US and Global Station Data**

Use Climvis to plot time-series of station climate data or contour/vectors map analyses for a region at one time. Station data is available globally w/some limitations and variables may include temperature (maximum, minimum and mean), winds, moisture, precipitation, pressure, snowfall, degree days, sunshine and visibility. Time-series plots for US climate division data (temperature, precipitation and Palmer Drought Index) and contour plots of GHCN temperature and precipitation are also available. From NOAA's National Climate Data Center (NCDC).

- **Western Regional Climate Center**

Western Regional climate center station data plots. Plot "normals" for station data in the western US. Variables include precipitation and temperature and statistics include monthly and daily averages, rainfall probabilities and related quantities.

Climate Datasets

- **Climate Reanalyzer**

Climate Reanalyzer is being developed by the Climate Change Institute at the University of Maine to provide an intuitive platform for visualizing a variety of weather and climate datasets and models. Investigate climate using interfaces for reanalysis and historical station data. Plot maps, timeseries, and correlations; export timeseries data to a text file for later use in spreadsheet software; export map layers to Google Earth.

- **NOMADS Climate Model Output:GFDL**

Users can plot Climate Model variables: Currently they have Dec-Cen and ocean assimilation datasets.

- **TAO Buoy Data**

Users can plot TAO buoy data.

- **Live Access to various climate data servers**

Users can plot climate data using the Live-Access-Server (LAS) from various collections. The Live Access Server (LAS) is a highly configurable web server designed to provide flexible access to geo-referenced scientific data. It can present distributed data sets as a unified virtual data base through the use of DODS networking. Ferret is the default visualization application used by LAS, though other applications (Matlab, IDL, GrADS, ...) can also be used.

- **Get and Plot IRI/LDEO Datasets**

Users can plot various datasets. Choose dataset of interest. The plotting options are at top of page that is returned.

- **EPIC**

The EPIC Web Database Web application that provides interactive access to EPIC hydrologic data sets. It allows users to view, download, and plot the data. The data sets include: data for the United States, data for the...



데이터 수집 및 저장측면

- 수문정보는 (국토교통부), 기상(기상청), 수질(환경부), 재난(국민안전처), 중소하천(지방자치단체), 다목적댐(한국수자원공사), 농업용저수지(한국농어촌공사) 등이 **각 기관의 목적에 맞게 각자 데이터베이스를 구축하여 운영**
- 각 기관은 타 기관과 협정을 통하여 데이터를 수집하여 각자 구축한 데이터베이스에 저장하고 있으나 **데이터 중복성 및 데이터 품질관리에 어려움**

데이터 공개 및 공동 활용측면

- 수요자의 needs 분석을 통한 정보의 공개 및 포털 구축 필요 - 주는 사람 마음대로?
- 다자간 데이터 연계는 일부 기능을 위한 데이터 연계에 불과하여 국가적인 차원에서 **데이터 연계 및 정보공동 활용을 통한 장기적인 대책** 시급 - **진정한 빅데이터 체제**

데이터 예측 및 운영 측면

- 타 기관에서 받은 데이터의 정확성이 확보되지 않은 상태에서 각 기관마다 확보한 데이터를 이용하여 **기관의 역할에 맞는 예측정보를 생산하기에는 어려움**
- 빅데이터 기반 수문기상관리체계 구축 및 장단기 예측을 통한 적극적인 운영이 필요함에 따라 **국가차원의 통합 수문기상자료 관리 시급**

4. Kwater 빅데이터기반 물정보관리

K-water 빅데이터 기반 물정보 통합관리체계 구축을 위한 마스터플랜 수립

K-water 사업분야별 업무 및 정보시스템 환경분석

물관련 기관 정보시스템 및 자료유통 현황분석

빅데이터 추진과제 도출 및 구체화

빅데이터 기반 K-water 물정보 통합관리체계 마스터플랜

Kwater는 우리나라에서 가장 다양한 업무를 기반으로 수문/환경/방재/기상 관련 빅데이터를 수집중

유역통합관리		수량관리		수질·수환경관리			시설안전관리	재난관리	
		유량관리(이수)	홍수관리(치수)	수질관리	생태환경관리	유사관리			지하수관리
주요 업무	본사	조시총괄, 이수안전평가 물 확보 등 이수사업 이수운영 기준수립·관리 유량분석·관리 기술개발	조시총괄, 치수안전평가 홍수방어 등 치수사업 홍수운영 기준수립·관리 홍수분석·관리 기술개발	조시총괄, 수질개선평가 탁수저감 등 수질개선사업 수질관리 기준수립·관리 수질예측 등 기술개발	조시총괄, 생태변화평가 서식처 등 생태관리사업 생태관리 기준수립·관리 생태서식해석 등 기술개발	조시총괄, 유사변화평가 유사조결 등 유사관리사업 유사관리 기준수립·관리 하상변동분석 등 기술개발	조시총괄, 지하수변화평가 지하수 보전·확보 등 사업 지하수관리 기준수립·관리 지하수해석 등 기술개발	조시총괄, 시설물 안전성평가 기초밀 성능·기능개선사업 지반관리 기준수립·관리 안전·유지관리 기준수립·관리	조시총괄, 재난대책평가 지반관리 기준수립·관리 지반대표표준매뉴얼관리 지반관리 등 기술개발
	수계	연간 이수운영계획 총괄 중·장기 기상분석·예보 발전사업·전력거래 총괄	연간 홍수운영계획 총괄 중·단기 기상분석·예보 발전사업·전력거래 총괄	수질관리 계획총괄 탁수관리 계획총괄 부유물관리 계획총괄	생태환경관리 계획총괄 퇴적도 오염관리 계획 총괄 환경영향평가 총괄	유사계획 총괄	지하수관리 기본계획수립	시설물 안전·유지분수계획총괄 안전관련 교육·훈련	기반상황·다응원화 총괄 (홍수·가뭄, 수질·시설안전사고 등)
연경	계측수립 조사분석 통합운영 유지관리 재난관리	수계 이수운영계획수립 유량분석, 모의운영 등 담·보·연계 공급량 결정 수계 발전시설통합운영 수계 수문·이수시설조사	수계 홍수운영계획수립 홍수분석, 모의운영 등 담·보 수문방류량 결정 수계 수문시설통합운영 수계 수문·치수시설조사	수계 수질관리계획수립 허천·지수지 수질수질분석·예측 오염원 관리, 탁수저감 수계 생태환경 조사 수계 퇴적도오염 관리 총괄	수계 생태환경관리 계획수립 생태환경관리 계획수립 수계 퇴적도오염 관리 총괄 수계 환경영향평가 총괄	퇴사, 유사량조사·분석 하상변동 조사 및 예측 준설 계획수립 및 시행	지하수 이용조사·분석 지하수 관정관리	담, 보, 제방 등 하천시설 정기·정밀점검, 안전점검 유지보수 계획수립·시행	지반다응원화매뉴얼관리 담·보 붕괴비상대책계획 상시 지반감시 지반상황 전파 및 대응 ·복구수행 수립 및 시행
	관계	수문·기상관측시설 관리 이수계약, 가용조사 지원	홍수예·경보시설 관리 수문방류시 통보·준설 수문조작(필요시)	수질 측정시설 관리 현장 수질조사 지원 부유물 수거·처리	에도, 생태하천 등 관리 퇴적도 오염 조사 및 관리 환경영향평가 조사 및 관리	유사, 하상변동조사 지원 준설 시행 관리	지하수 현장조사 지원	시설안전 계획기 관리 (토압, 수압, 침하계 등)	지반감시 (CCTV 등)관리 지반발생시 현장 대응
관계기관	기상청·국토부·환경부 지자체·한수원·농촌공	기상청·국토부(FCO) 소방청·지자체 등.	환경부(유역환경청)	국토부·환경부·지자체	국토부·지자체	국토부·농림부·지자체	국토부·소방청 시설안전공단	국토부·환경부·형안부 소방청·국방부·지자체	

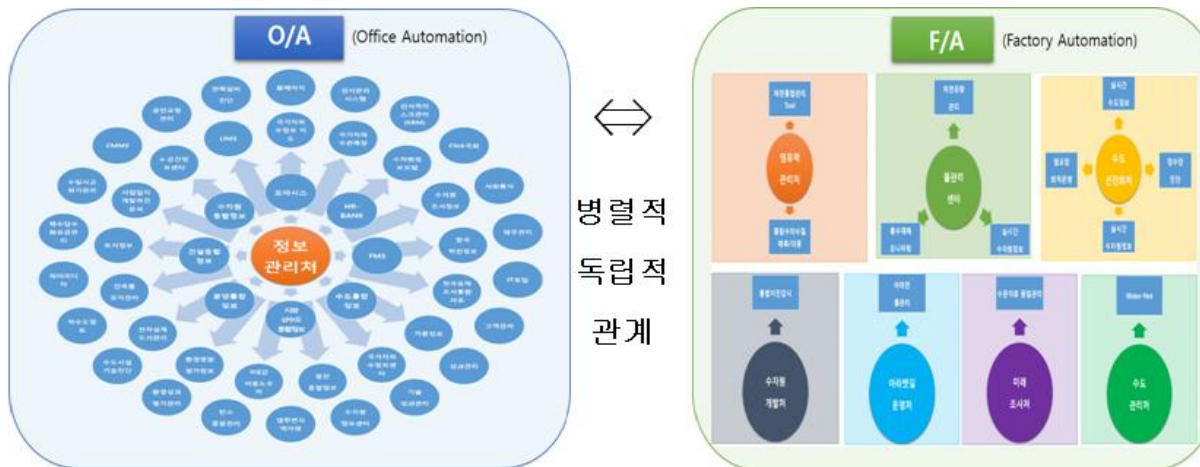


사업분야(수자원/수도/도시,환경)별 축적된 자료 관리를 위해 30여개 정보시스템 및 20여개 분석시스템 운영중 - Work-Flow/시스템/정형,비정형 데이터 분석

데이터	관측·측정 자료						영상·GIS 자료			
	기상관측	증발량	토양수분	우량·수위	유량유사량	수질지하수	위성영상	수치지형	하천지형	
분석시스템	조사		운영			관리				
	수문자료 품질관리		하천유량관리	K-HIT	홍수재해 모니터링	통합수리수질 예측 및 대응	하천통합관리 Tool	댐안전관리 통합	통합지진감시	
	<ul style="list-style-type: none"> 자료 비교분석 이상치 관리 오결측자료 검정 		<ul style="list-style-type: none"> 강우-유출모형 저수지 연계모형 댐-보 연계운영 하천물수지분석 수질모형 	<ul style="list-style-type: none"> 강우예측 실시간수문정보 통합수문자료 물재해감시 저수지용수공급 홍수분석 발전통합운영 	<ul style="list-style-type: none"> 재난상황 통합 관리 및 제어 기관 자료연계 도달시간 분석 <p>* 지자체 해외사업 대상 패키지SW</p>	<ul style="list-style-type: none"> 유역단위 감시·모니터링 수리수질예측 오염원 추적 	<ul style="list-style-type: none"> 수위-수질 분석 유역유출 분석 오염사고 분석 3D GIS 시각화 4대강 리버뷰 	<ul style="list-style-type: none"> 위험인자 해석 위험도 분석평가 	<ul style="list-style-type: none"> 시설물 모니터링 지진계 관리 	
정보시스템	조사					운영·관리				
	한국하천정보 시스템	수자원조사 정보시스템	유역조사통합 자료조회	가뭄정보 시스템	국가지하수 정보센터	수자원통합 정보시스템	실시간수자원 정보	아라천물관리 시스템		
	<ul style="list-style-type: none"> 하천통합검색 하천영상 하천역사, 법령 하천지도, 유역도 	<ul style="list-style-type: none"> 유량성과 관리 수문조사 교육 수문조사연보 	<ul style="list-style-type: none"> 유역조사성과 관리 및 조회 (유역, 시계열, 행정구역별 조회) 	<ul style="list-style-type: none"> 가뭄현황 및 전망 (기상학적가뭄 모니터링) 	<ul style="list-style-type: none"> 지하수관측연보 이용량 통계 지하수관리계획 개발절차 지하수디지털맵 	<ul style="list-style-type: none"> 경영정보 댐유역기상자료 수문정보 댐저수지 수질 댐용수영업 댐안전관리 발전종합 	<ul style="list-style-type: none"> 댐-보 수문정보 CCTV 영상 GIS기반 하천정보 예경보시설 관리 수문관측소 관리 	<ul style="list-style-type: none"> 아라천 기상수문 수질, 유수소통 선박운항 시설물 정보 		

경영 등 일반 업무프로세스 시스템(OA, 정보관리처 주관) 45개와 계측 Data 감시 및 제어시스템(FA, 사업부서 주관) 13개 구축 운영 중

<p>(경영지원) HR-Bank, 재무통합(FMS), OASIS, KRM, 법무, 감사, 고객관리 등 총 17개 시스템 운영</p>	<p>(수자원) 수자원통합(WRIS), 수자원조사, 하천, 가뭄, 지하수정보, 발전통합 등 총 18개 시스템 운영</p>	<p>(수도) 수도통합(WIS), 지방상수도통합(Water-INFO), 시설물유지관리(CMMS), 수운영시스템(iWater) 등 총 16개 시스템 운영</p>	<p>(건설관리·도시환경 등) 건설통합(CMS), 분양통합정보, 전자설계도서, 토지정보, 건축물유지관리 등 총 7개 시스템 운영</p>
--	--	---	--



(OA : 정보관리처 주관)

(FA : 사업부서 주관)

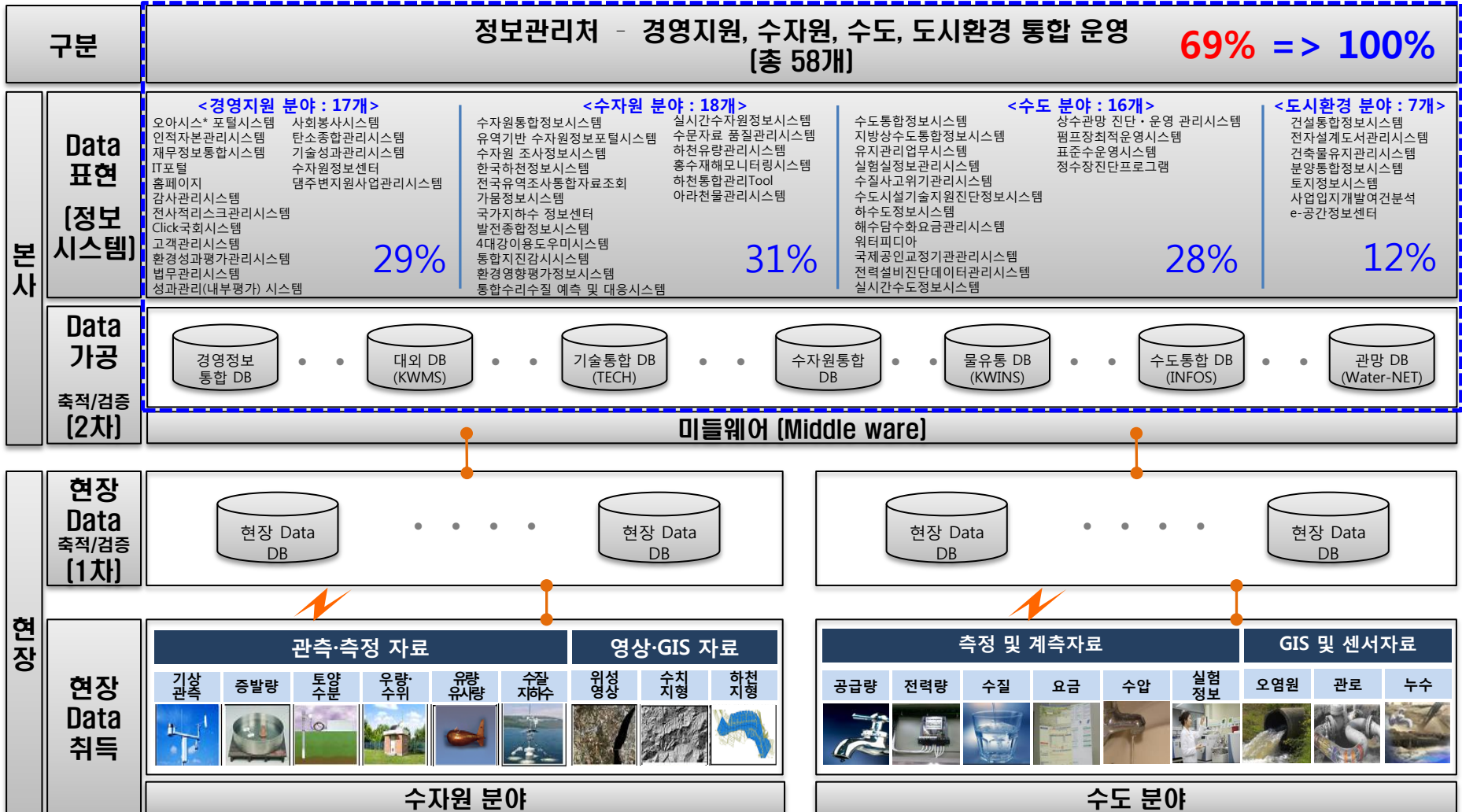
2014년 기준

구 분		정보시스템 현황
OA (45)	경영 (17)	홈페이지, 감사관리시스템, 전사적리스크관리시스템, Click국회, 사회봉사시스템, 법무관리시스템, IT포털, 고객관리시스템, 성과관리시스템, 기술성과관리시스템, 수자원정보센터, 댐주변지원사업관리시스템, 탄소종합관리시스템, 환경성과평가관리시스템, 오아시스, 인적자원관리시스템, 재무통합정보시스템
	수자원 (10)	수자원정보포털시스템, 수자원조사정보시스템, 한국하천정보시스템, 가뭄정보시스템, 전국유역조사통합자료조회, 국가지하수정보센터, 발전종합정보시스템, 4대강이용도우미시스템, 환경영향평가정보시스템, 수자원통합정보시스템
	수도 (11)	수도시설기술지원진단정보시스템, 하수도정보시스템, 해수담수화요금관리시스템, 워터피디아, 수질사고위기관리시스템, 유지관리업무시스템, 국제공인교정관리시스템, 전력설비진단데이터관리, 실험실정보관리시스템, 수도통합정보시스템, 지방상수도통합정보,
	건설관리 (7)	전자설계도서관리, 건축물유지관리, 토지정보, 사업입지개발여건분석, e-공간정보센터, 건설통합정보, 분양통합정보
FA (13)	수자원 (8)	하천통합관리Tool, 통합수리·수질예측및대응시스템, 하천유량관리시스템, 홍수재해모니터링시스템, 실시간수자원정보시스템, 통합지진감시시스템, 아라천물관리시스템, 수문자료품질관리시스템
	수도 (5)	실시간수도정보시스템, 펌프장최적운영시스템, 수운영시스템, 정수장진단프로그램, 상수관망진단운영관리시스템

사업분야별(경영지원, 수자원, 수도 등) 정보 인프라/시스템의 독립화

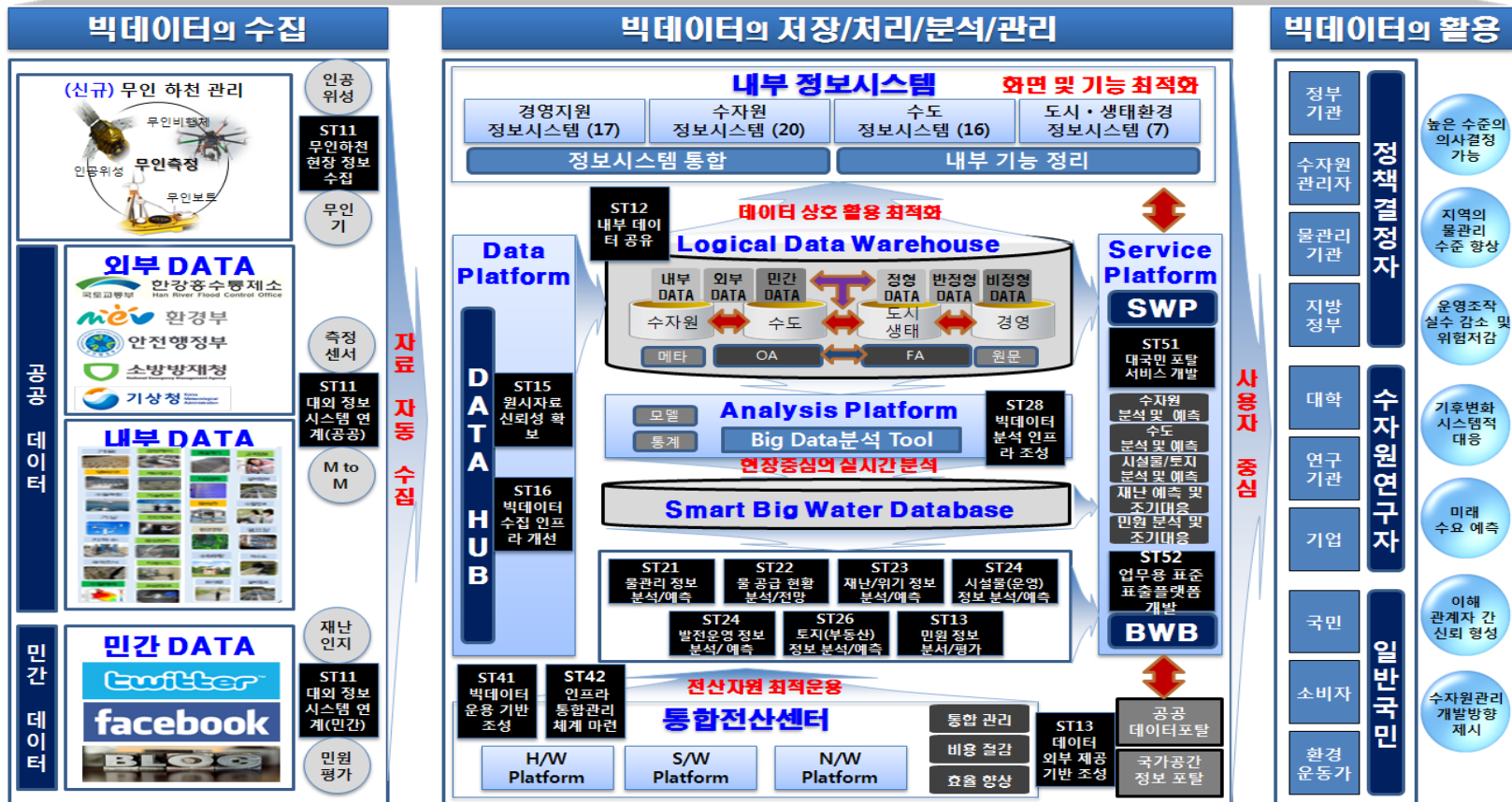


사업분야별(경영지원, 수자원, 수도 등) 정보 인프라/시스템의 통합화

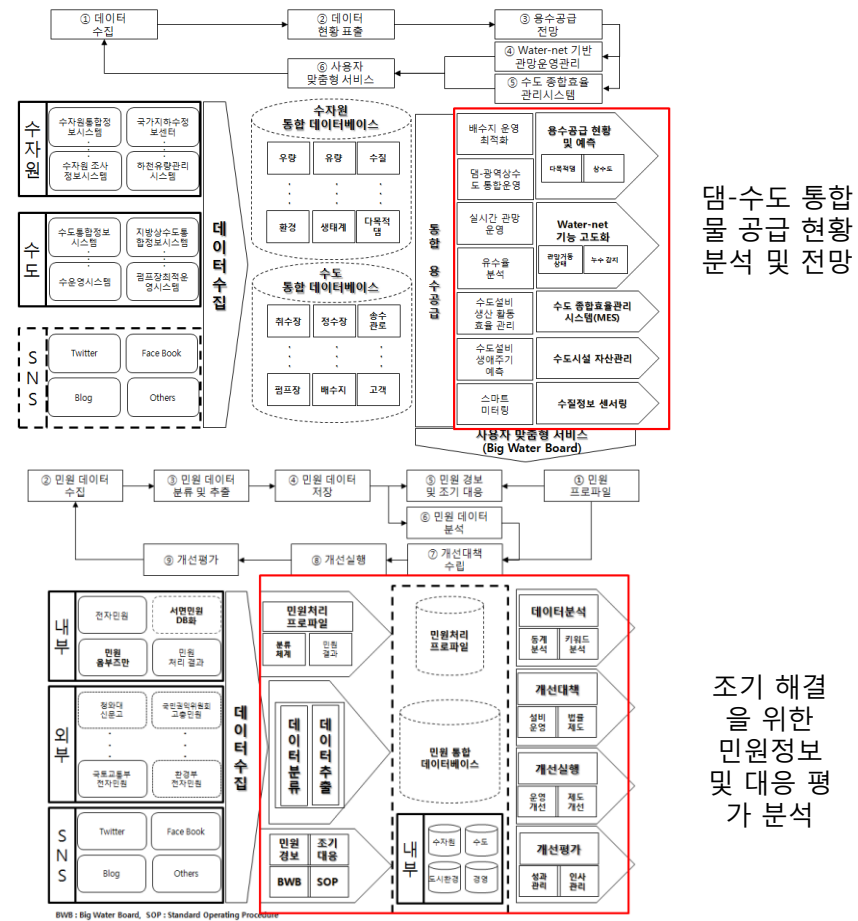


향후 빅데이터 추진은 빅데이터 솔루션보다 필요 데이터 확보 및 확보한 데이터의 분석 및 활용이 중요

K-water 빅데이터 기반 물정보 통합 플랫폼 개념도



진정한 빅데이터를 위해서는 현장/외부 데이터 확보/연계활용/데이터 오픈 등이 관건



조기경보 및 대응을 위한 재난/위기 정보 분석 및 예측

조기 해결을 위한 민원정보 및 대응 평가 분석

K-water Big-Data 기반 물정보 통합관리체계 개념도



5. 우리의 과제

빅데이터 도구 사용법을 익히는 데는 시간과 노력이 필요합니다.

- 목수가 가구를 만들때 사용하는 연장에 익숙해져야 하는 것과 같이
- 화가가 그림을 그리는 캔버스와 붓과 물감에 익숙해져야 하는 것과 같이
- 빅데이터 도구도 익숙해지는데 숙련이 필요합니다.

빅데이터 도입하기

- 학생들에게 빅데이터 도구를 사용할 수 있는 기회를 많이 갖도록 합니다.
- 빅데이터 처리에 효과적인 R, Matlab, Python이나 Scalar 같은 프로그래밍 언어를 습득하도록 합니다.
- 여러 프로젝트에 빅데이터 도입을 적극적으로 반영하여 ICT 전문가들과 교류하도록 합니다.
- 수문해석 기법을 빅데이터 방식으로 접근하는 것에 대해 많이 생각해보도록 합니다.








빅데이터의 세계는 신대륙과 같습니다.

- 빅데이터를 도입하는 것은 신대륙으로 향하는 여행과 같은 것으로 약간의 위험을 감수한다면
- 분명 새로움과 즐거움이 우리를 기다리고 있을 것입니다.
- 제자를 가르치면서, 친구와 같이한다면 이 여행도 그리 나쁘진 않을 것입니다.

연구목표

- 빅데이터 기반 가뭄관리를 위한 **수문기상정보 활용기술** 제시
- 빅데이터 기반의 **가뭄 감시/피해 및 영향 분석/전망 기술** 개발

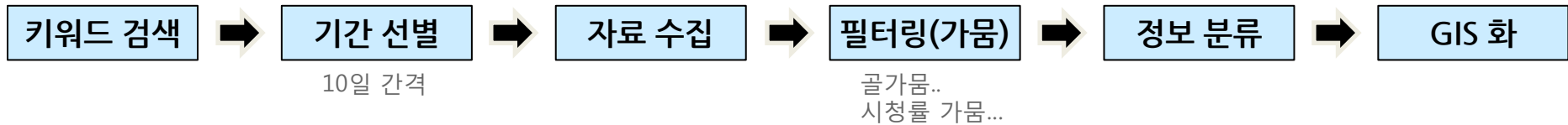
빅데이터 기술 연구

구분	설명
	<ul style="list-style-type: none"> • 기반 시스템(HW/건물)은 클라우드 IaaS 서비스를 제공하는 KT cloud를 사용하고 있습니다. • 클라우드 서비스는 컴퓨팅 자원을 빌려쓰리는 것으로 CPU, Memory, Storage, Network를 별도 구입이나 구축없이 그냥 전기나 수도처럼 필요시 계약하고 사용한 만큼 비용을 지불하는 것입니다. • 빅데이터를 사용하면 엄청난 자원을 사용하므로, 기존 컴퓨터를 사용하면 비효율적입니다.
	<ul style="list-style-type: none"> • 운영 시스템(OS/운영자)은 리눅스의 일종인 Ubuntu를 사용했습니다.- 빅데이터는 리눅스 친화적입니다. • Ubuntu는 윈도우 만큼이나 사용하기 편리합니다.
	<ul style="list-style-type: none"> • 빅데이터 기반 구축(Platform/식당)은 Hadoop(MapReduce/HDFS)을 사용하여 구축했습니다. • Hadoop은 여러 대의 컴퓨터를 이용할 수 있는 기반을 제공합니다.
	<ul style="list-style-type: none"> • NoSQL(분산형 DB/주방)의 구축은 Hbase를 사용하여 구축했습니다. • NoSQL이 오라클과 구분되는 기능입니다. Hbase는 데이터를 key-value 형식으로 관리합니다.
	<ul style="list-style-type: none"> • Opentsdb(맛있는 양념)는 시계열 데이터를 NoSQL의 key-value 타입으로 구축할 수 있도록 지원합니다. • Opentsdb는 Hbase로의 입력을 담당합니다.
	<ul style="list-style-type: none"> • 자료의 통계처리(자료분석/요리방법)를 위해서는 R 을 사용했습니다. • The R Project for Statistical Computing
	<ul style="list-style-type: none"> • 빅데이터 가시화 플랫폼 (자료출력/서빙)은 Granfa입니다. • 현재 기본 세팅만 되어 있는 상태입니다.

자료 입력 순서 : Json(데이터) → Script(작성코드) → Opentsdb → Hbase → Hadoop
 자료 출력 순서 : Grafana ← Opentsdb ← Hbase ← Hadoop

◆ 가뭄피해 및 가뭄영향 분석 및 정보 제공

▪ 빅데이터 수집 절차



▪ 가뭄피해 및 영향 관련 빅데이터 분류 키워드 (안)

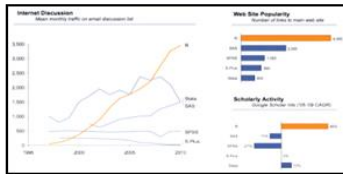
농업	산업	수생태계	사회, 공중보건	물공급	대응, 제한
곡물가격상승 대체작물재배 수확량감소 소득감소 파종 모내기지연 작물고사 농작물 피해 농작물재해보험	용수부족 공장가동중단	물고기폐사 용존산소량 염도상승 녹조 조류농도	갈수현상 수질저하 수인성 전염병	비상식수 급수지원 식수공급 급수차	재난관리기금 재난지역 지원 예비비 관정개발 긴급자금

◆ 빅데이터 기반 가뭄감시

국가가뭄정보 포털 Korean Drought Portal

기상청

- 기상학적 가뭄
- 유역별 강수량 상황



Kwater

- 수문학적 가뭄
- 생공용수 물부족 지도
- 동네 물수급 상황

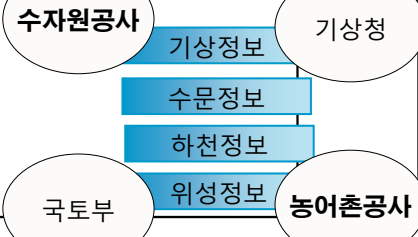
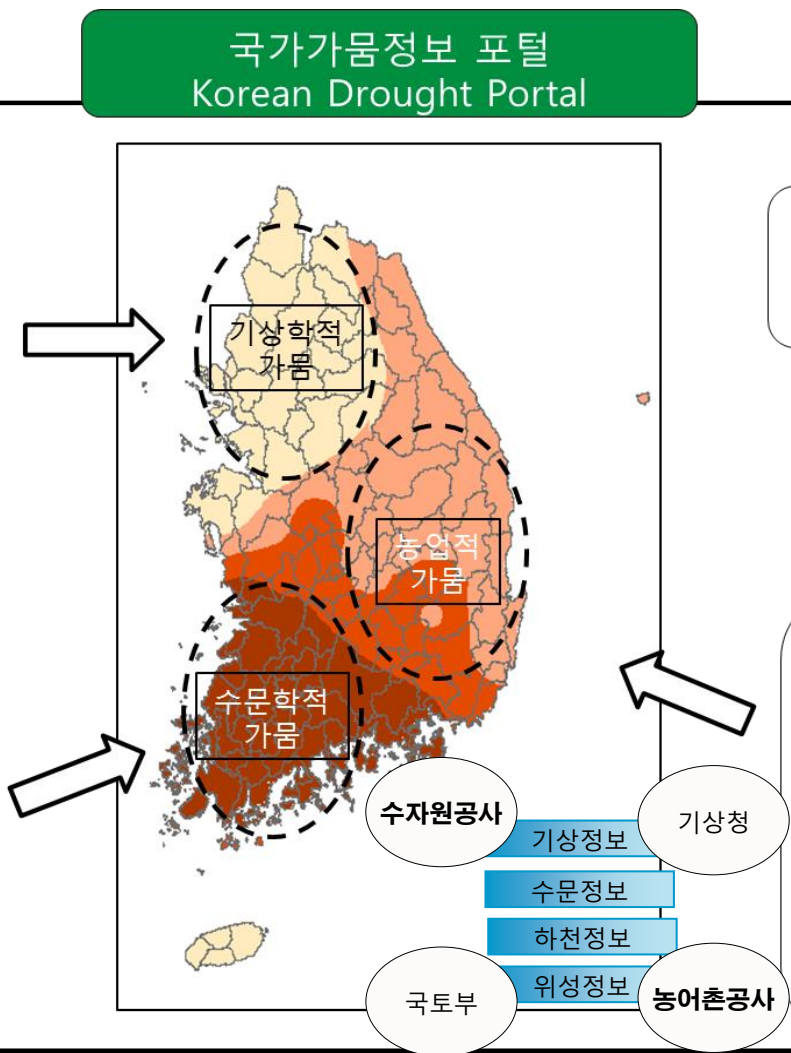
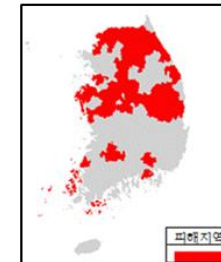


국민안전처

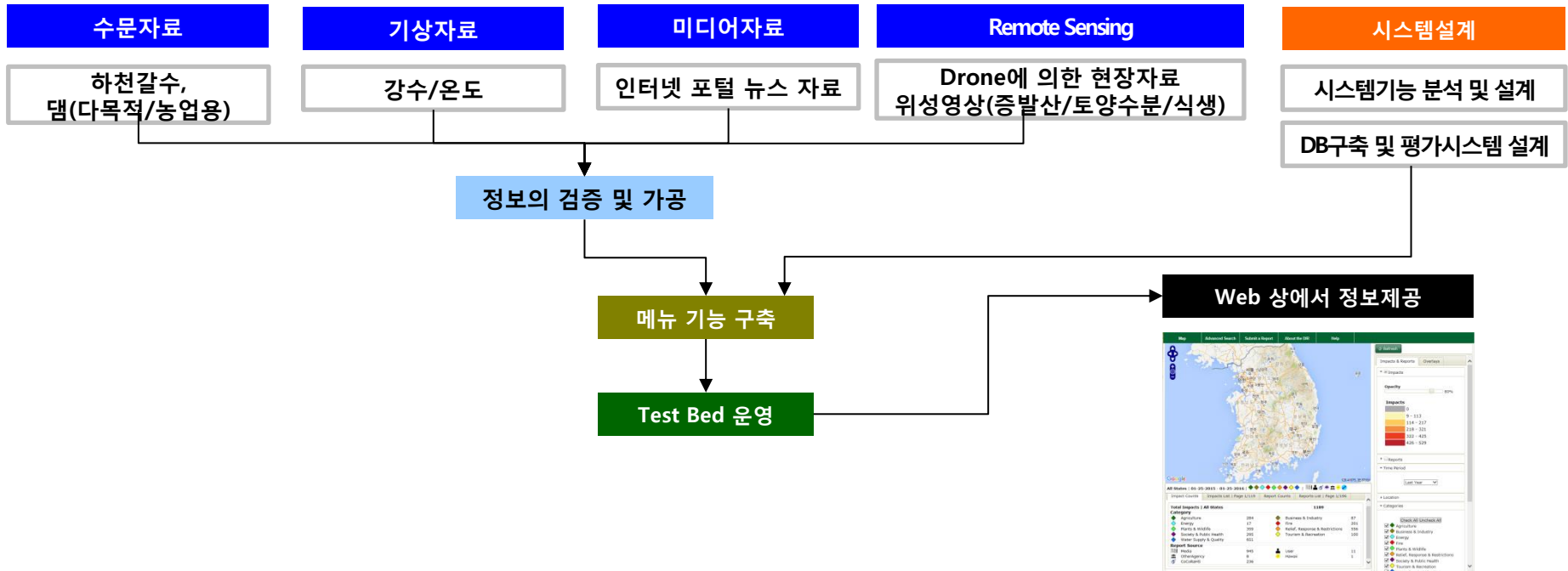
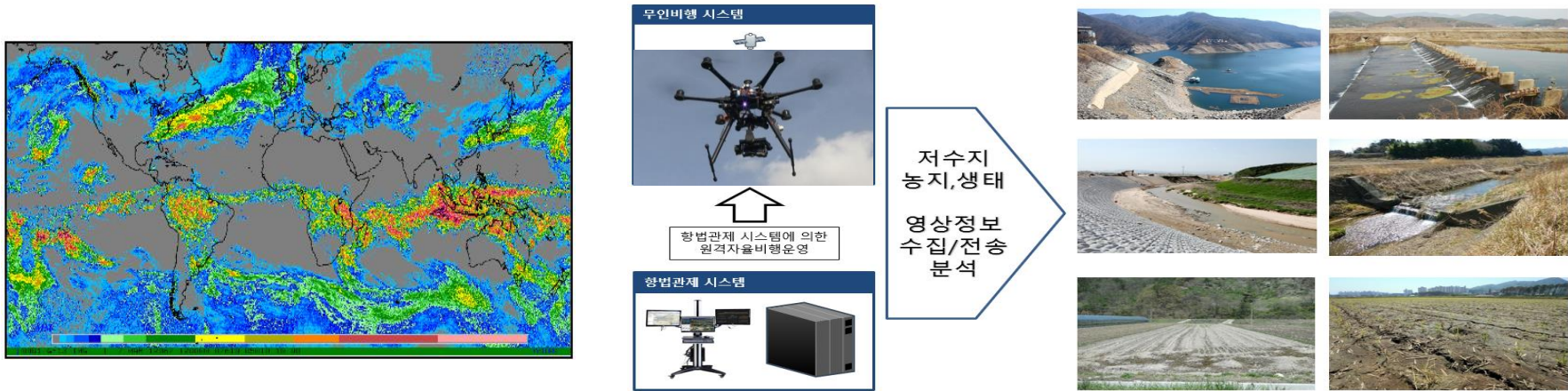
- 가뭄피해현황
- 가뭄피해액 집계
- 가뭄대책의 추진

농촌공사

- 농업적 가뭄
- 농작물 피해 현황
- 관개용수 부족도
- 토양수분도



우리팀은 빅데이터를 이렇게 활용합니다.



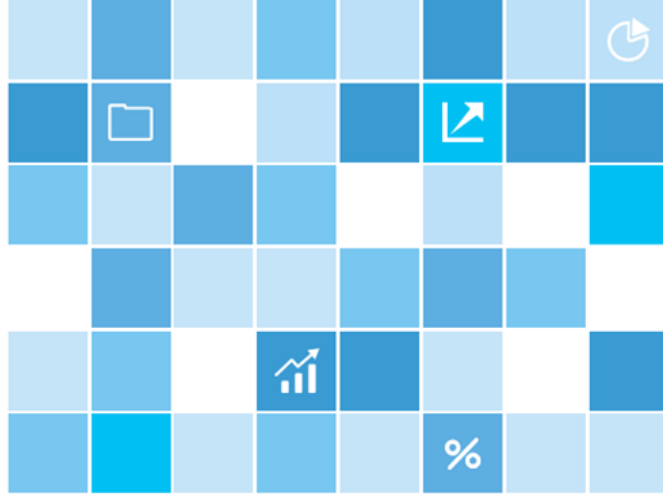


- 현재 Hadoop → Hbase → OpenTSDB를 이용하여 NoSQL 시스템을 구축하고 Grafana를 이용하여 시각화 시스템을 구축하였습니다.

빅데이터의 미래와 우리의 현재

- 우리는 정형데이터에 집착
- 한가지 목적을 달성하기 위한 시스템 개발 (특정 부처의 업무목적 달성을 위한 시스템)에 집착
- 다른 분야와의 공유에 인색 / 정보(자료공개) 공개에 인색
- 주는 사람 마음대로 / 받는 사람은 ?
- **생각을 크게, 유연하게, 모두가 함께 !!!**
- **또 다른 차원의 홍수/가뭄 예측**
- **기후변화 연구의 새로운 패러다임**
- **Big data open platform for efficient water resources management**
- **수도/수문(지표수,지하수)/기상(기후)/수질/위성 자료의 공동 활용**





Thank You

이주현

leejh@joongbu.ac.kr | +82-31-8075-1615