

what the heck is

데이터 거버넌스

목차

- 데이터 거버넌스란?
- 누가 해야 하나?
- 추진경과 : 우물 밖 vs. 우물 안
- 데이터 혁신 로드맵 수립 프로젝트 (PwC 이성근 이사)

데이터

- 연설의 뒤에 있는 진실
- 보고의 뒷배
- “최적의 LP호가 최소 스프레드는”
 - 설문조사? 공청회?
 - 누구를 위한 최적인가?

데이터 거버넌스란?

















000 총류

010 도서학, 서지학
020 문헌정보학
030 백과사전
040 강연집, 수필집, 연설문집
050 일반연속간행물
060 일반학회, 단체, 협회, 기관
070 신문, 언론, 저널리즘
080 일반저집, 총서
090 향토자료

100 철학

110 형이상학
120
130 철학의 세계
140 경학
150 아시아철학, 사상
160 서양철학
170 논리학
180 심리학
190 윤리학, 도덕철학

200 종교

210 비교종교
220 불교
230 기독교
240 도교
250 천도교
260 신도
270 바라문교, 인도교
280 회교
290 기타 제종교

300 사회과학

310 통계학
320 경제학
330 사회학, 사회문제
340 정치학
350 행정학
360 법학
370 교육학
380 풍속, 민속학
390 국방, 군사학

400 순수과학

410 수학
420 물리학
430 화학
440 천문학
450 지학
460 광물학
470 생명과학
480 식물학
490 동물학

500 기술과학

510 의학
520 농업, 농학
530 공학, 공업일반
540 건축공학
550 기계공학
560 전기공학, 전자공학
570 화학공학
580 제조업
590 가정학 및 가정생활

600 예술

610 건축술
620 조각
630 공예, 장식미술
640 서예
650 채화, 도화
660 사진술
670 음악
680 연극
690 오락, 운동

700 언어

710 한국어
720 중국어
730 일본어
740 영어
750 독일어
760 프랑스어
770 스페인어
780 이탈리아어
790 기타 제어

800 문학

810 한국문학
820 중국문학
830 일본문학
840 영미문학
850 독일문학
860 프랑스문학
870 스페인문학
880 이탈리아문학
890 기타 제문학

900 역사

910 아시아
920 유럽
930 아프리카
940 북아메리카
950 남아메리카
960 오세아니아
970 양극지방
980 지리
990 전기

현대소설

813.6



주류
(문학)



강목
(한국)

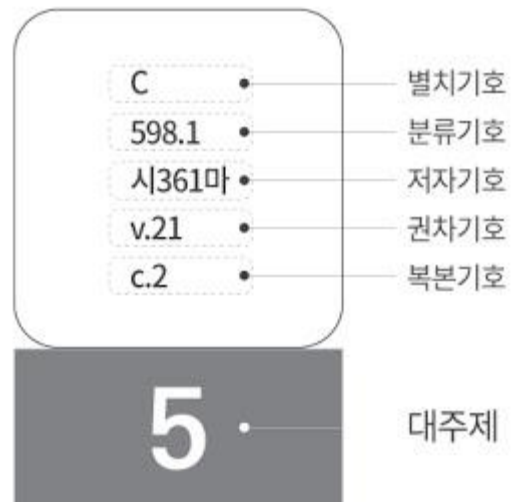


요목
(소설)



세목
(현대)

청구기호라벨



1. 별치기호 - 소장된 위치를 구분하기 위해 표시
2. 분류기호 - 한국십진분류법으로 분류한 번호
3. 저자기호 - 저자명과 도서명을 조합한 기호
4. 권차기호 - 전집이나 시리즈 책번호 구분
5. 복본기호 - 동일한 책이 2권이상 일때 표시
6. 대주제 - 분류기호에 따른 색라벨

〈표 1〉 미·한·일 도서 십진분류법 비교

번호	듀이십진분류법	한국십진분류법	일본십진분류법
000	총류	총류	총기(總記)
100	철학	철학	철학
200	종교	종교	역사
300	사회학	사회학	사회과학
400	언어	자연과학	자연과학
500	자연과학	기술과학	기술·공학·공업
600	기술과학	예술	산업
700	예술	언어	예술
800	문학	문학	언어
900	역사	역사	문학

〈표 2〉 교보문고 도서분류체계

기호	분류
A	잡지, 만화
B	여행, 취미, 건강, 예술
C	일본서적
D	외국어
E	외국서적
F	유아, 아동
G	중·고학습
H	경제, 경영, 자기계발, 정치, 사회, 수험서
I	과학, 기술, 공학, 컴퓨터, 멀티미디어, 책공방
J	소설, 시, 에세이
K	인문, 역사, 문화, 종교

데이터 거버넌스란?

How do you define data governance?



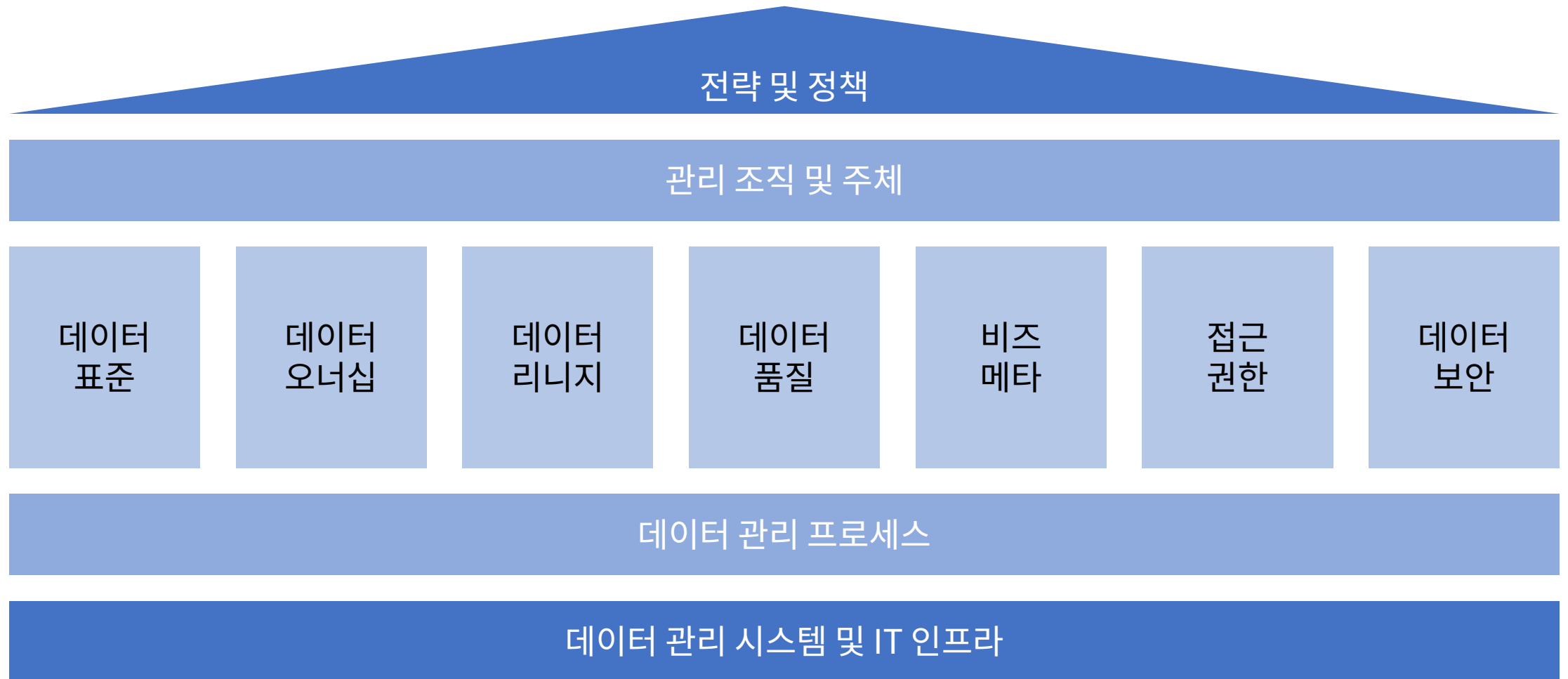
Note: Multiple responses allowed

Data: UBM survey of 118 business technology professionals at organizations with 1,000 or more employees, November 2017

“기업이 데이터를 잘 활용하기 위해서
무엇을 어떻게 해야 하는가?”

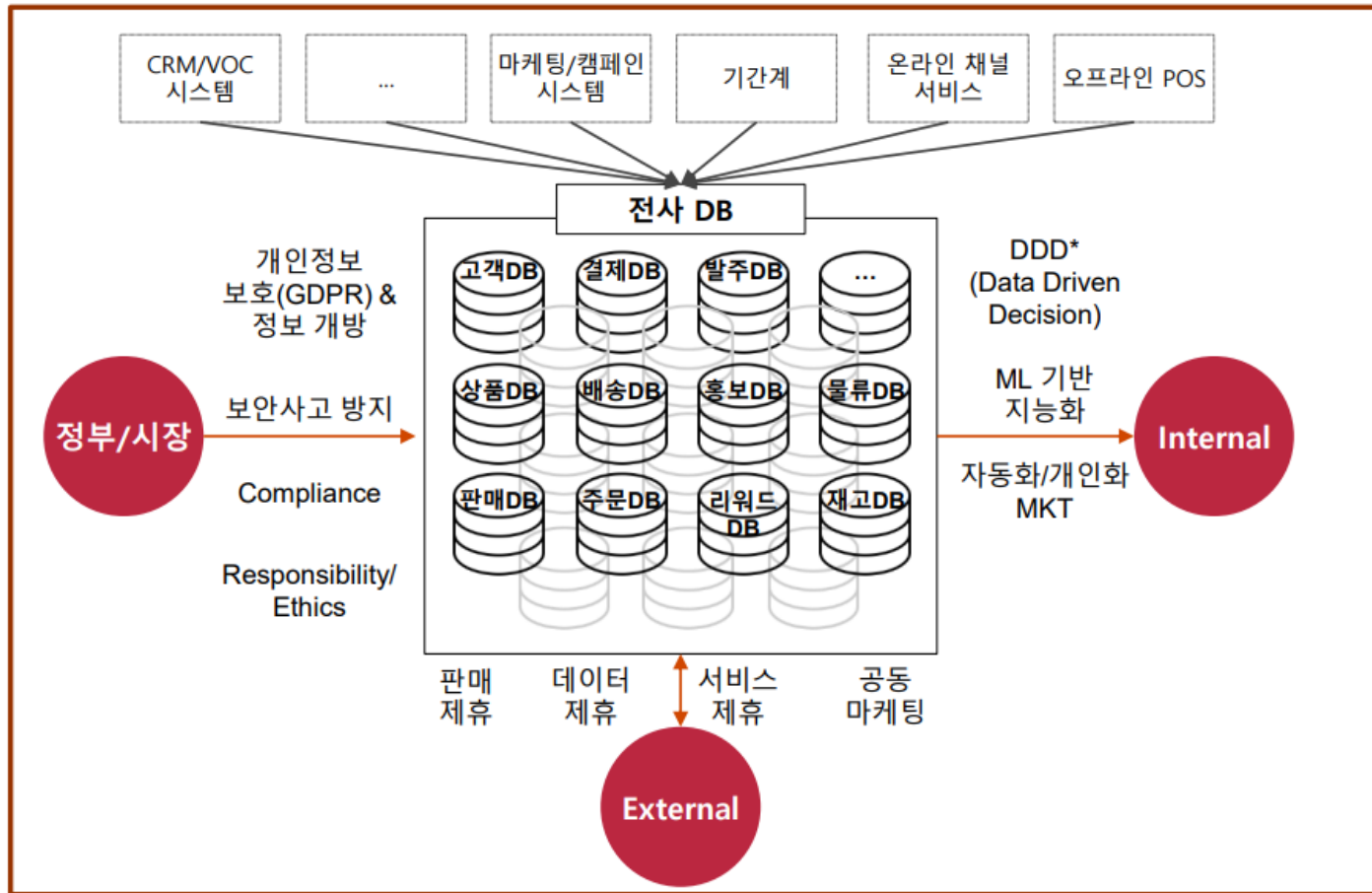
에 대한 답을 위해 필요한 모든 활동·체계

데이터 거버넌스의 구성



데이터 거버넌스 꼭 해야하나?

“다 모르지만, 한 가지는 확실하게 답할 수 있다”



데이터는
정확한가요?

OOO는 무슨
데이터인가요?

XXX 데이터
가지고 있나요?

규정 개정의
영향도 분석은
금방 되나요?

글쎄요...
확인해보겠습니다.

글쎄요...
확인해보겠습니다.

글쎄요...
확인해보겠습니다.

놉!!

누구에게 물어봐야 하나요?

Why does it matter?



"Please ask the table owner"



"I'm no longer responsible for this table, please ask team X"



"This is an upstream problem, we can't fix it"



"How do I find the pipeline owner?"



IT에 물어보세요!!

...

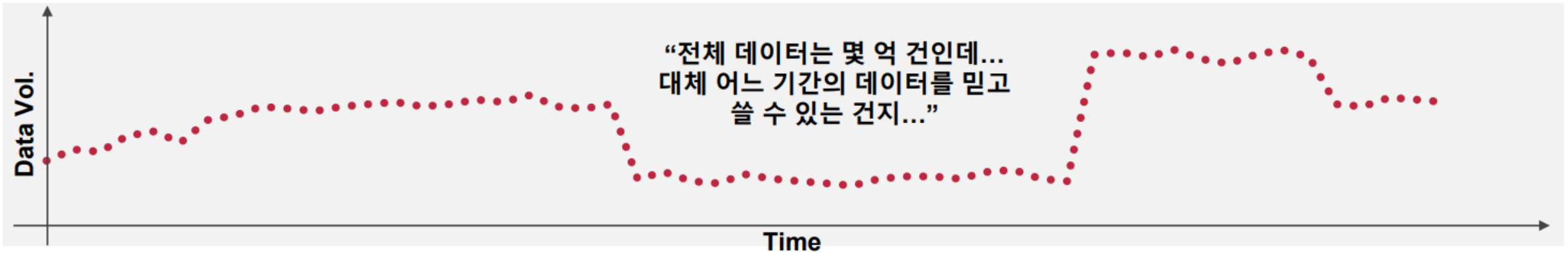
오늘 휴가입니다.!



Multiple days

데이터가 있는데, 없습니다

- 거래소 수수료 수입 추이를 보고하시오.



- 데이터가 잘못됐나? 데이터에 도대체 뭘 짓을 한거야?!
- 히스토리를 아는 사람만 제대로 보고할 수 있다!

데이터 거버넌스가 없으니...

- 데이터가 있는데, 쓸 수 없다
- 데이터가 있는지조차 모른다
- 여러 부서에 동일한 데이터가 각각 쌓인다
- 누가 데이터 오너인지 모른다
- 왜 사용할 수 없는지 모른다
- 사용을 위해 누구와 협의해야 하는지 모른다

데이터 가공 능력이 없으면 벌어지는 일

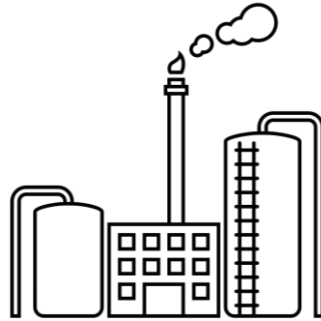
- 데이터 가공 기관으로부터 인덱스 상품 구성 정보 구매를 고려 중
- 정제기술이 없는 후진 산유국이 원유를 팔고 정제유를 수입하는 셈



매매정보



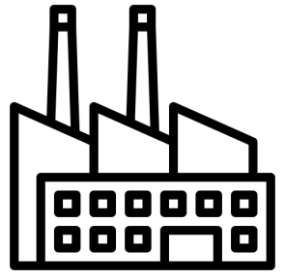
정보분배



데이터 기관



인터넷



투자자
금융기관
감독기관



기업정보

데이터 표준이 없으면 벌어지는 일

깊이 있는 통계

기업분석보고서 · 투자분석정보 · ESG 정보 · 이용안내

KOR | ENG KRX 정보데이터시스템

가격정보

거래정보

추세정보

기업정보

참고정보

수익률 상관관계

투자분석정보 / 가격정보 / 수익률 상관관계

상위종목

차트

종목비교

종목명

005930/삼성전자

Q

비교종목

A066570/LG전자

Q

+

조회

전체조건 닫기

종목명

005930/삼성전자

Q

비교종목

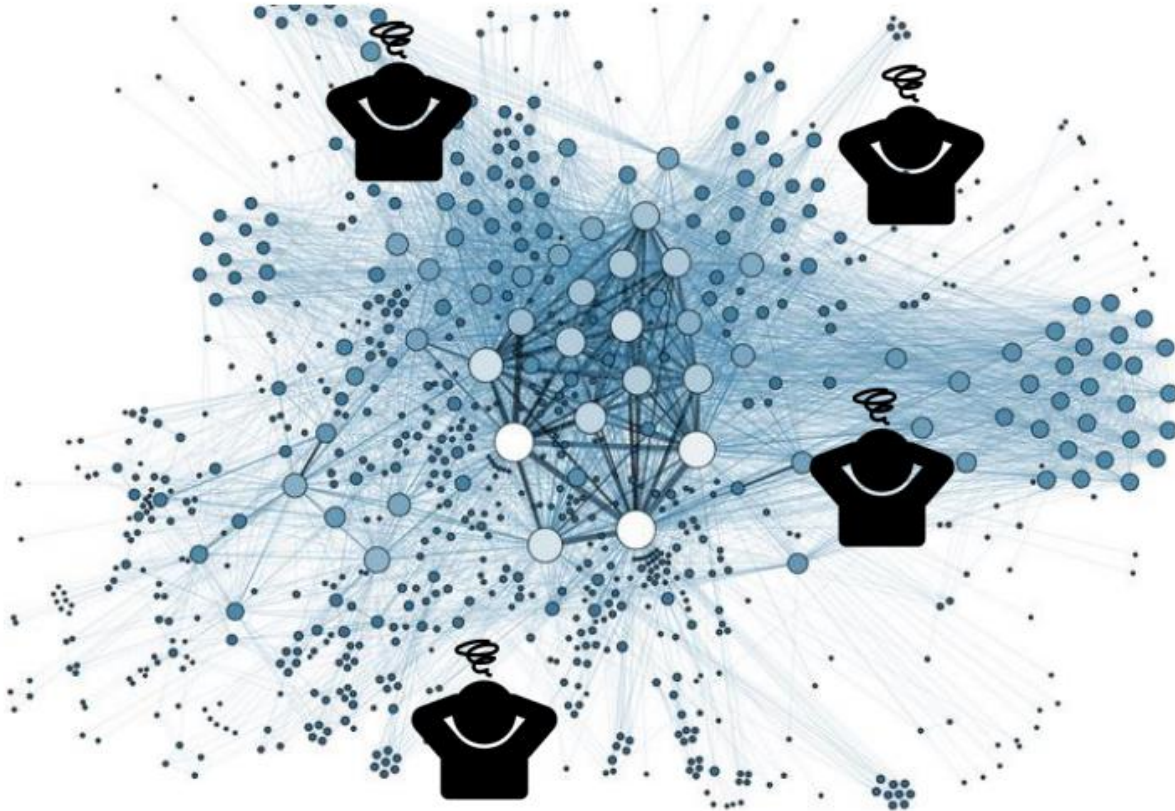
A066570/LG전자

Q

+

비교종목은 10개까지 추가 가능합니다. 비교종목 검색 후 추가해주세요.

이런 문제가 발생한 진짜 이유!



데이터가 너무 많아져서, 오히려 역설적으로 시기
적절하게 데이터를 활용하기 어려워졌다

- **데이터가 너무 많음**
 - 비슷한 이름의 Table, Column이 수십 개
 - 제대로 쌓이고 있는 건지 ...
- **데이터가 너무 빨리 변함**
 - 시장의 변화가 빨라지면서, 새로 쌓고, 변경되고, 중단하고
 - 질문/관점이 바뀌면서 필요한 데이터가 계속 변화
- **비즈니스 질문 복잡도 증가**
 - 단순 정보는 볼만큼 봤는데, 이제는 연결해야 하는데 ...
 - 데이터 **생산자 / 가공자 / 사용자** 모두 달라지고 있음

누가 해야 하나?

영어로?

정의

영어로?

지수

제일 먼저 떠오르는 것?

삼대

1,000,000

— 1,000,000 —

백만

100만

One million

1000000



현업



이 데이터는,

- 누가?
- 언제?
- 어디서?
- 어떻게?
- 왜?

만드나?

생산, 관리를 위한 표준
오너십, 권한관리

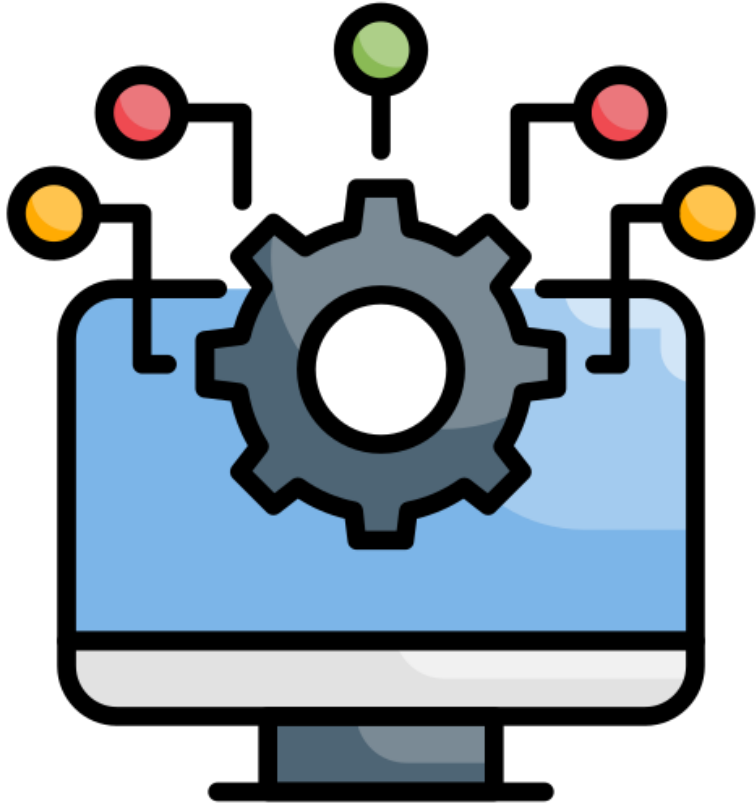
이 데이터를,

- 누가?
- 언제?
- 어디서?
- 어떻게?
- 왜?

사용하나?

분석, 활용을 위한 표준

IT



- 데이터 정합성 관리
- 성능 관리
- 보안 관리
- 시스템 도입, 구축, 운영

비즈메타

메타

- 일반적 의미
 - 그리스어 meta에서 유래, "초월", "넘어서", "뒤에"를 의미
 - 주로 어떤 개념이나 체계를 넘어서 그것을 성찰하거나 분석하는 과정, 또는 그 위에 위치하는 개념
- 메타데이터(Metadata) - 데이터의 데이터
 - 데이터를 설명하는 정보.
 - 사진 파일의 메타데이터: 촬영 날짜, 위치 정보, 파일 크기



12/24 ?



12/24 ?

12월 24일

0.5

“12/24”



12/24 ?

<날짜> 12월 24일 </날짜>

<수식> 0.5 </수식>

<문자> “12/24” </문자>

가보자고!

Before

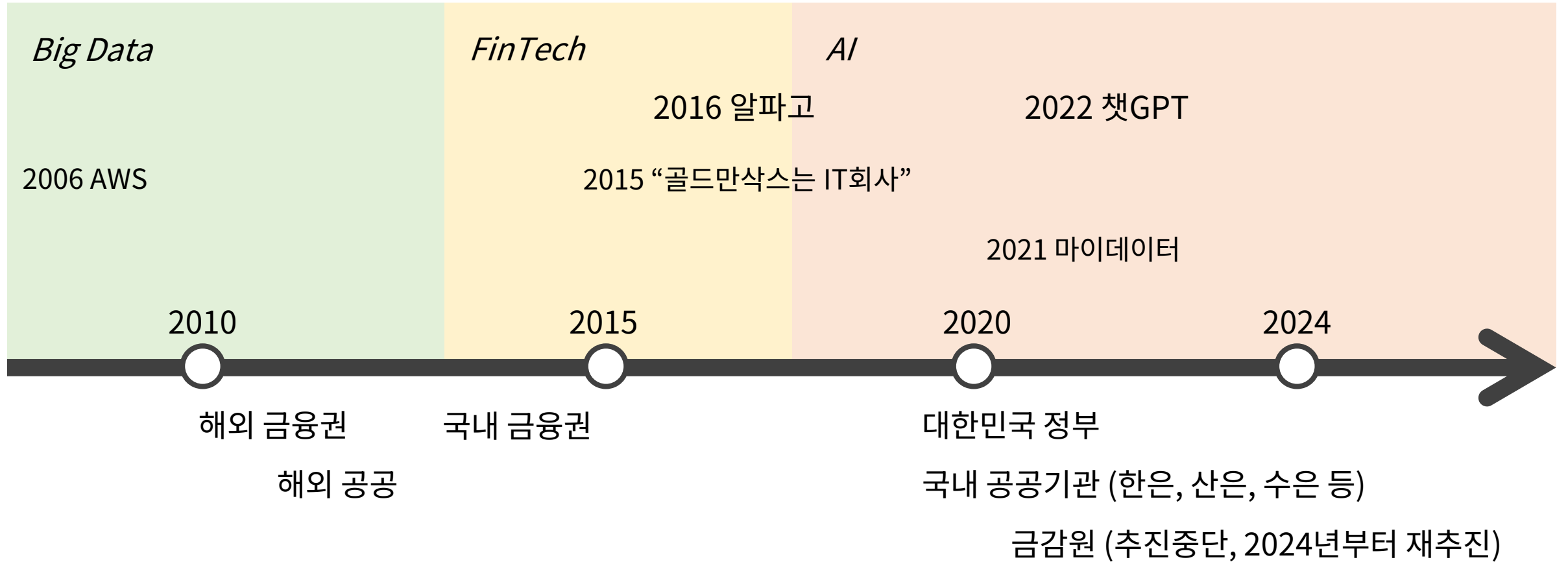


After



추진경과 : 우물 밖 vs. 우물 안

우리만 모르는 세상



우리가 아는 세상

