

Module 14. 프로젝트 방법론

02

빅데이터 플랫폼 구축

경북대학교 배준현 교수
(joonion@knu.ac.kr)

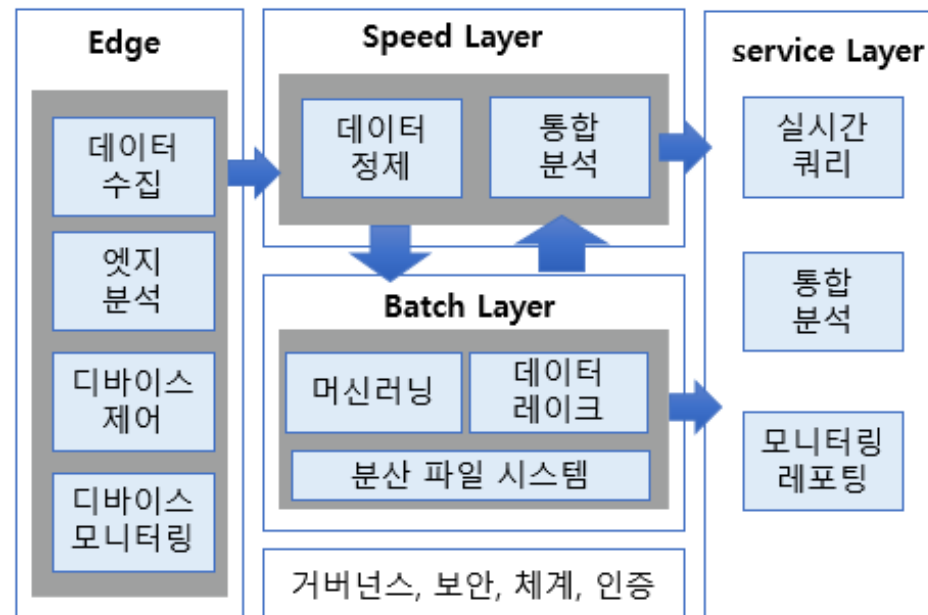


01. 데이터 분석 프로젝트

■ 빅데이터 플랫폼

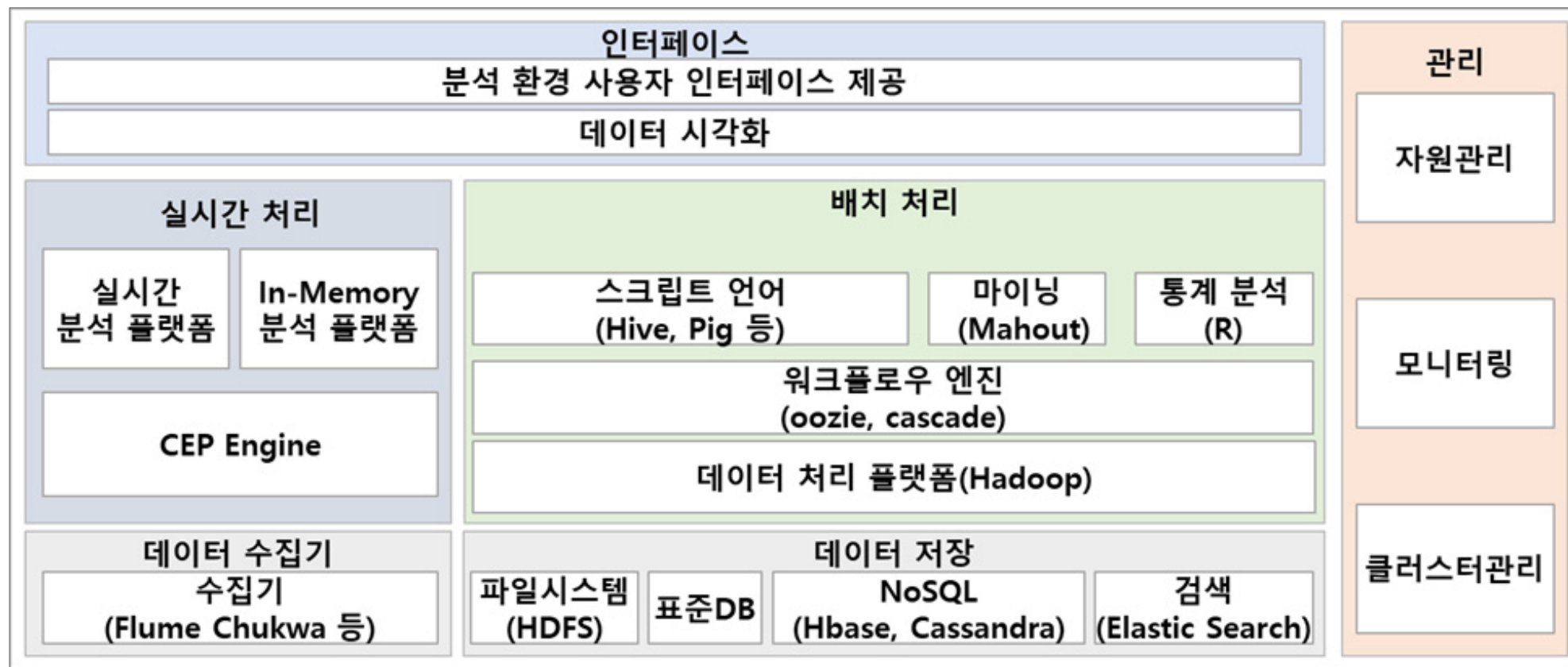
• 빅데이터 플랫폼은

- 다양한 소스로부터 생성되는 대량의 데이터는 처리하기 위하여
 - 소스: Computer, Internet, Mobile, IoT
- 데이터의 수집, 저장, 처리, 분석, 시각화를 제공한다.





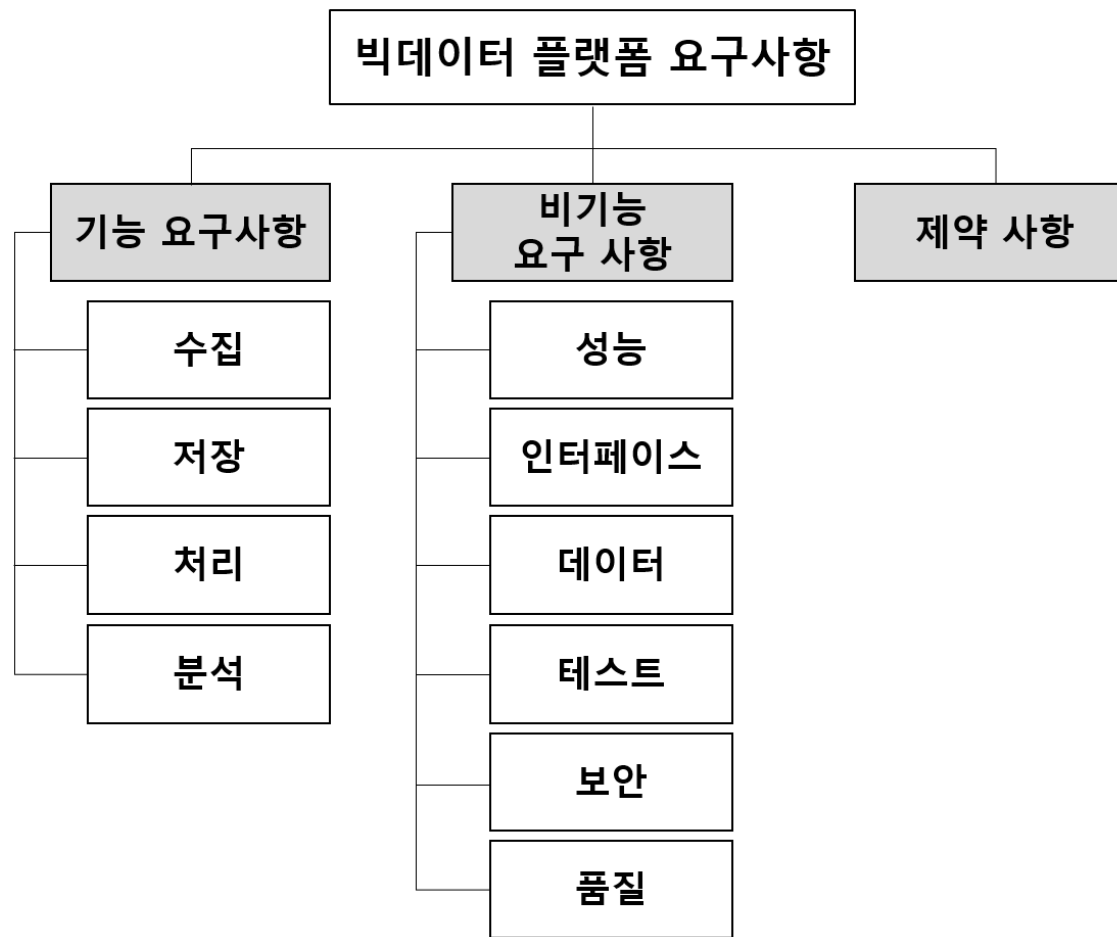
01. 데이터 분석 프로젝트





01. 데이터 분석 프로젝트

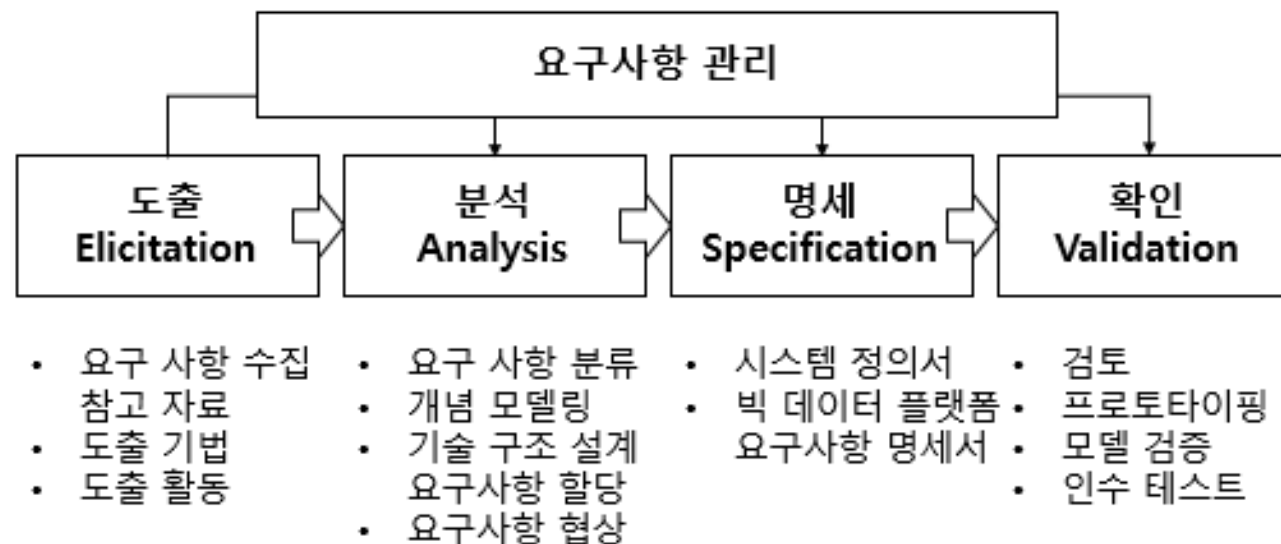
■ 빅데이터 플랫폼의 요구사항





01. 데이터 분석 프로젝트

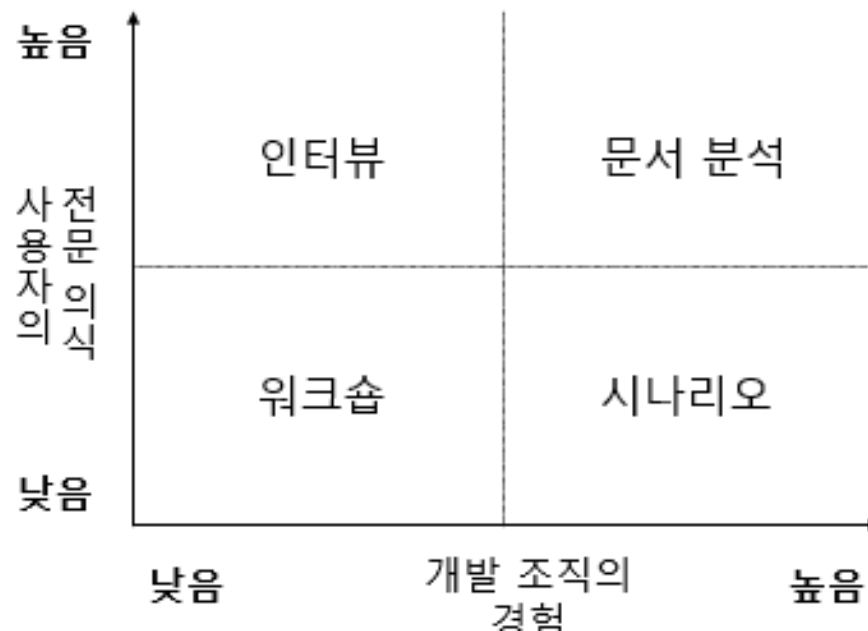
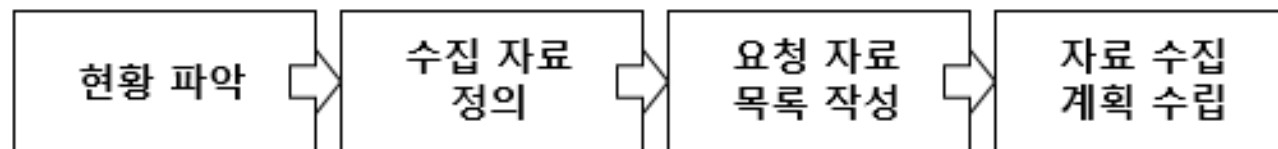
■ 요구사항 개발 프로세스





01. 데이터 분석 프로젝트

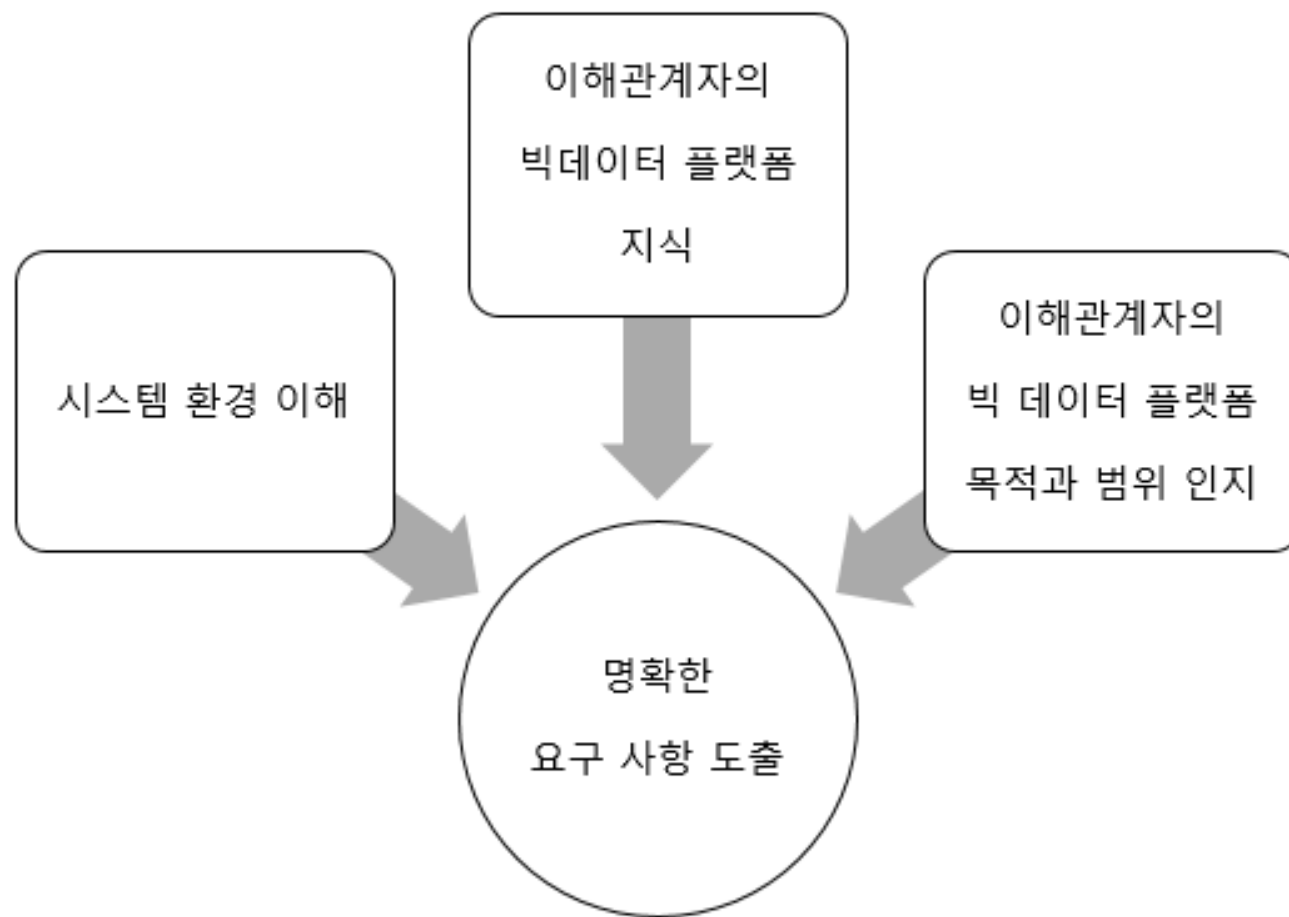
■ 요구사항 수집을 위한 사전 준비와 수집 기법의 활용





01. 데이터 분석 프로젝트

■ 이해관계자의 지식과 이해를 반영한 요구사항의 도출





01. 데이터 분석 프로젝트

■ 빅데이터 플랫폼의 요구사항 분류 기준

- 요구사항은 기능인가, 비기능인가?
- 요구사항이 하나 이상의 고수준 요구사항으로부터 유도된 것인가?
 - 또는 이해관계자나 다른 소스로부터 직접 발생한 것인가?
- 요구사항이 제품에 관한 것인가 프로세스에 관한 것인가/
- 요구사항의 우선순위는 어떻게 되는가?
- 요구사항이 소프트웨어 생명 주기 동안 변경이 발생하는가?



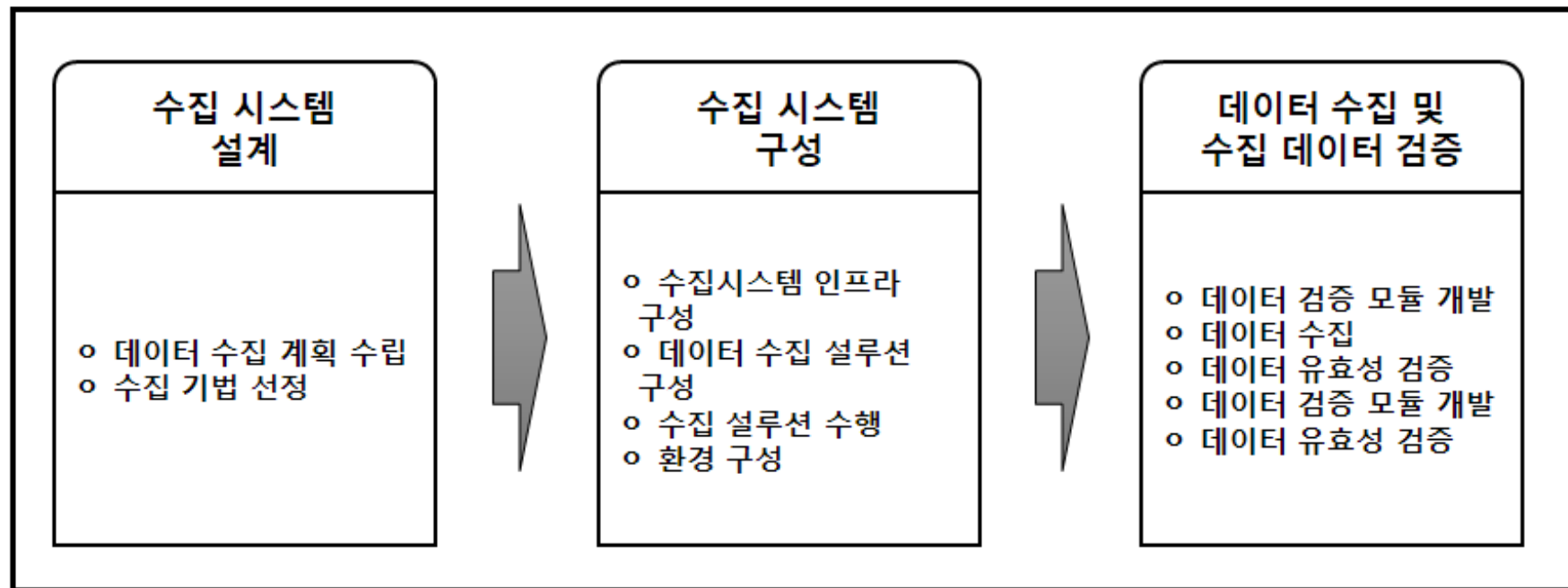
01. 데이터 분석 프로젝트

구 분	내 용
비즈니스 요구 사항	재무, 시장, 고객이나 개발 조직이 얻고자 하는 비즈니스 혜택이다. 예: - “A 지역의 시장 점유율을 3개월 안에 25% 향상시켜야 합니다.”
사용자 요구 사항	사용자 목표나 사용자가 수행해야 하는 비즈니스 업무를 표현한 일반적인 문장이며, 일반적으로 유스케이스, 시나리오, 사용자 스토리보드로 표현한다. 예: - “저는 선임 장비 관리자로서 매일 아침 펌프 컨트롤러를 조정해야 합니다.”
비즈니스 규칙	고객이 특정 조건하에서 어떤 활동을 수행해야 하는 것을 말한다. 예: - “반드시 준수해야 한다.” - “반드시 ~에 따라야 한다.” 등
기능적 요구 사항	특정 조건 하에서 시스템이 보여 주거나 사용자가 시스템을 사용할 수 있게 하는 행동을 식별할 수 있게 하는 행위를 말한다. 예: - “사용자는 프로젝트 목록을 알파벳순이나 역순으로 정렬할 수 있어야 합니다.”
품질 속성	시스템이 무엇을 얼마나 잘 제공하느냐에 대한 설명으로 빠르게, 쉽게, 사용자 친화적인, 신뢰할 수 있는, 안전한, 바람직한 등의 단어를 활용한다. 예: - “모바일 소프트웨어는 사용자의 터치 명령에 빠르게 반응해야 합니다
외부 인터페이스 요구 사항	시스템 자체를 제외한 모든 것의 연결하는 단어로, ‘~로부터 신호를 읽어야 한다.’, ‘사용자 인터페이스 요소는 <표준>을 준수해야 한다.’ 등과 같은 문구이다. 예: - “모바일 앱(app)은 제가 송금하는 수표를 촬영하고 나면 그 이미지를 은행에 전송할 수 있어야 합니다.”



01. 데이터 분석 프로젝트

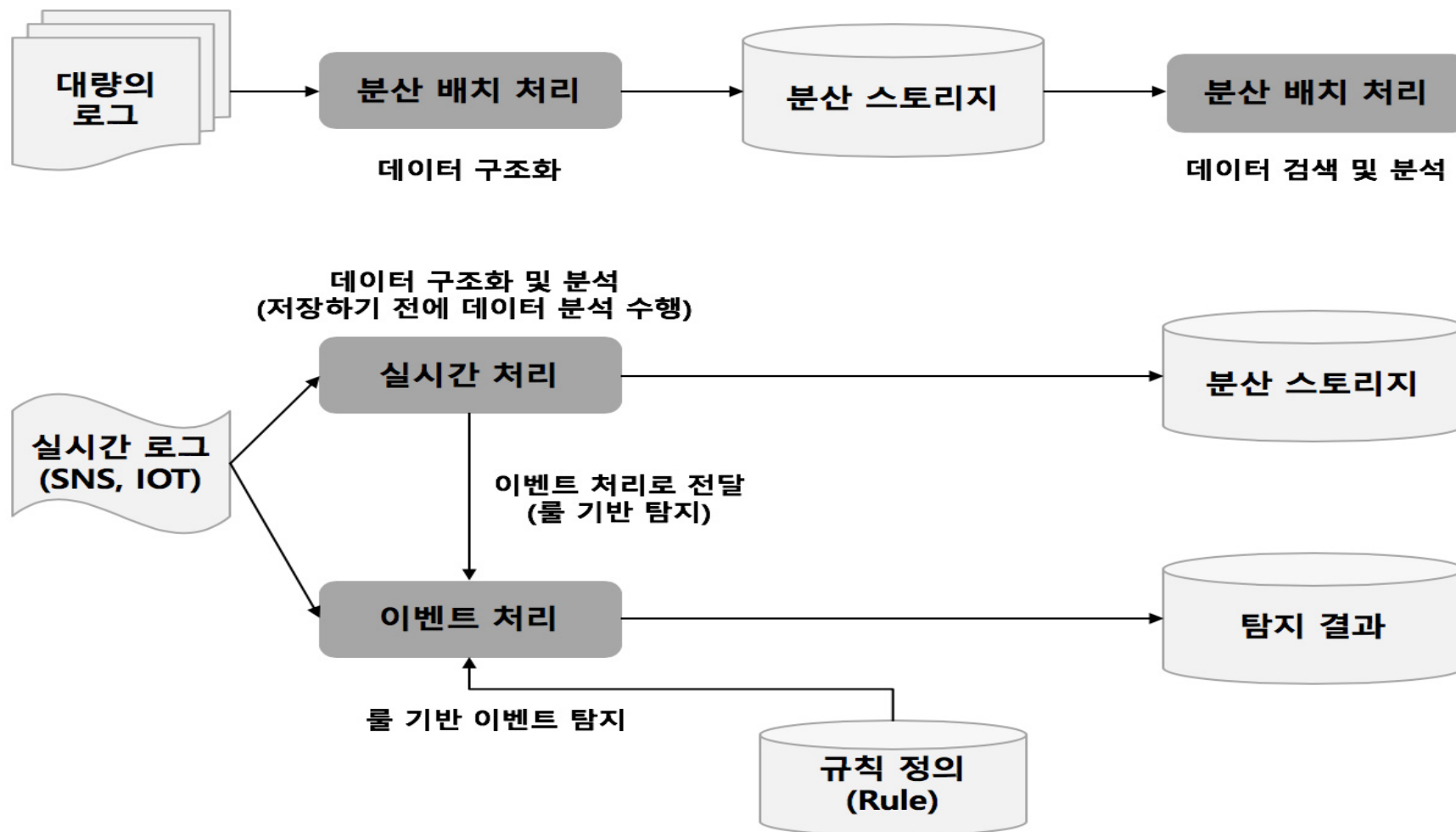
■ 빅데이터 수집 시스템 설계





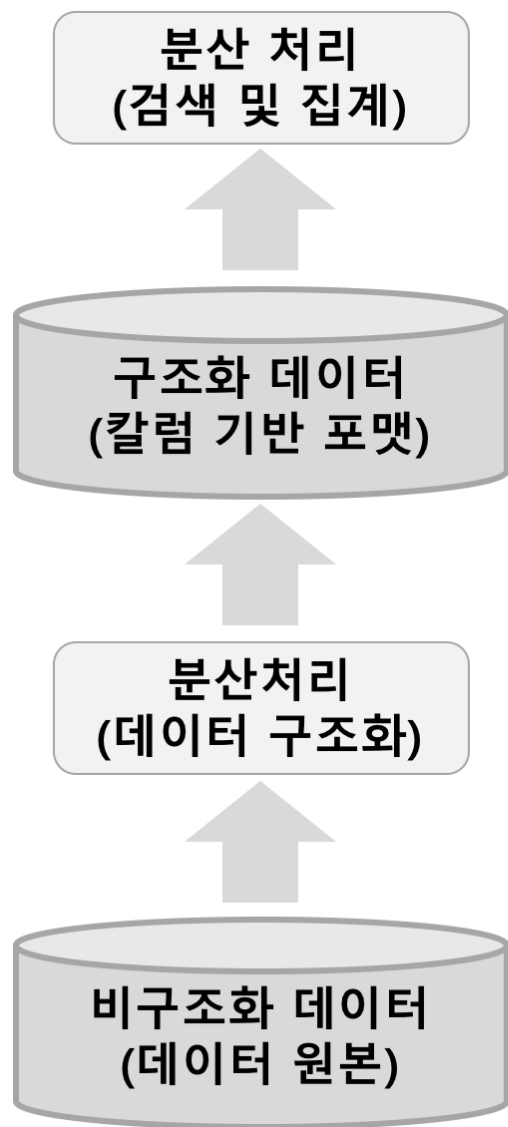
01. 데이터 분석 프로젝트

■ 빅데이터 처리 시스템 설계





01. 데이터 분석 프로젝트



- 대량의 데이터 검색을 위한 분산처리
- 복잡한 대량 데이터 분석 용도 (Hive)
- 빠른 데이터 검색 용도 (Impala, Presto 등)

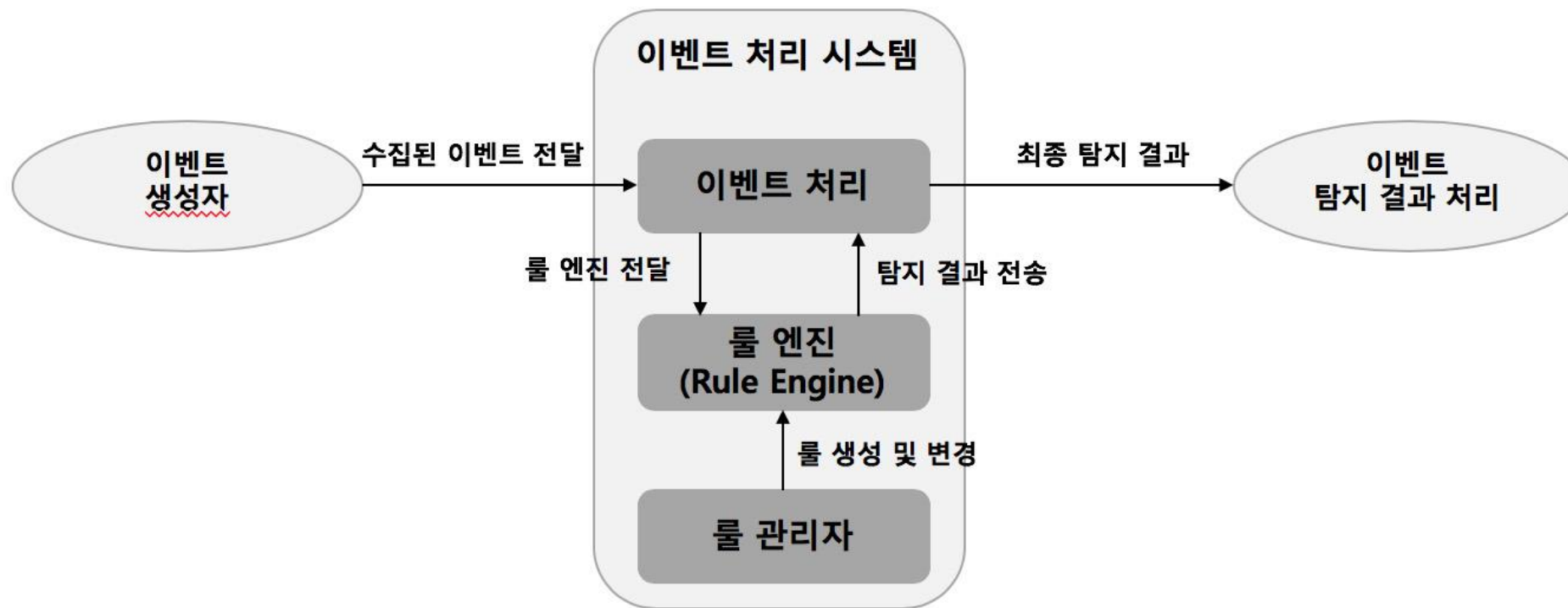
- 구조화된 데이터 (데이터 스키마 존재)
- 칼럼 기반 스토리지 적용
- 데이터 용량 및 검색에 최적화된 저장 구조

- 원본에서 데이터 가공 및 구조화를 위한 분산처리
- 저장 시 검색에 최적화된 압축 포맷으로 변경
- MapReduce, Apache Spark

- 수집된 원본 상태
- 별도 스키마가 없음. (문자열)
- 데이터 압축 적용 안 됨.

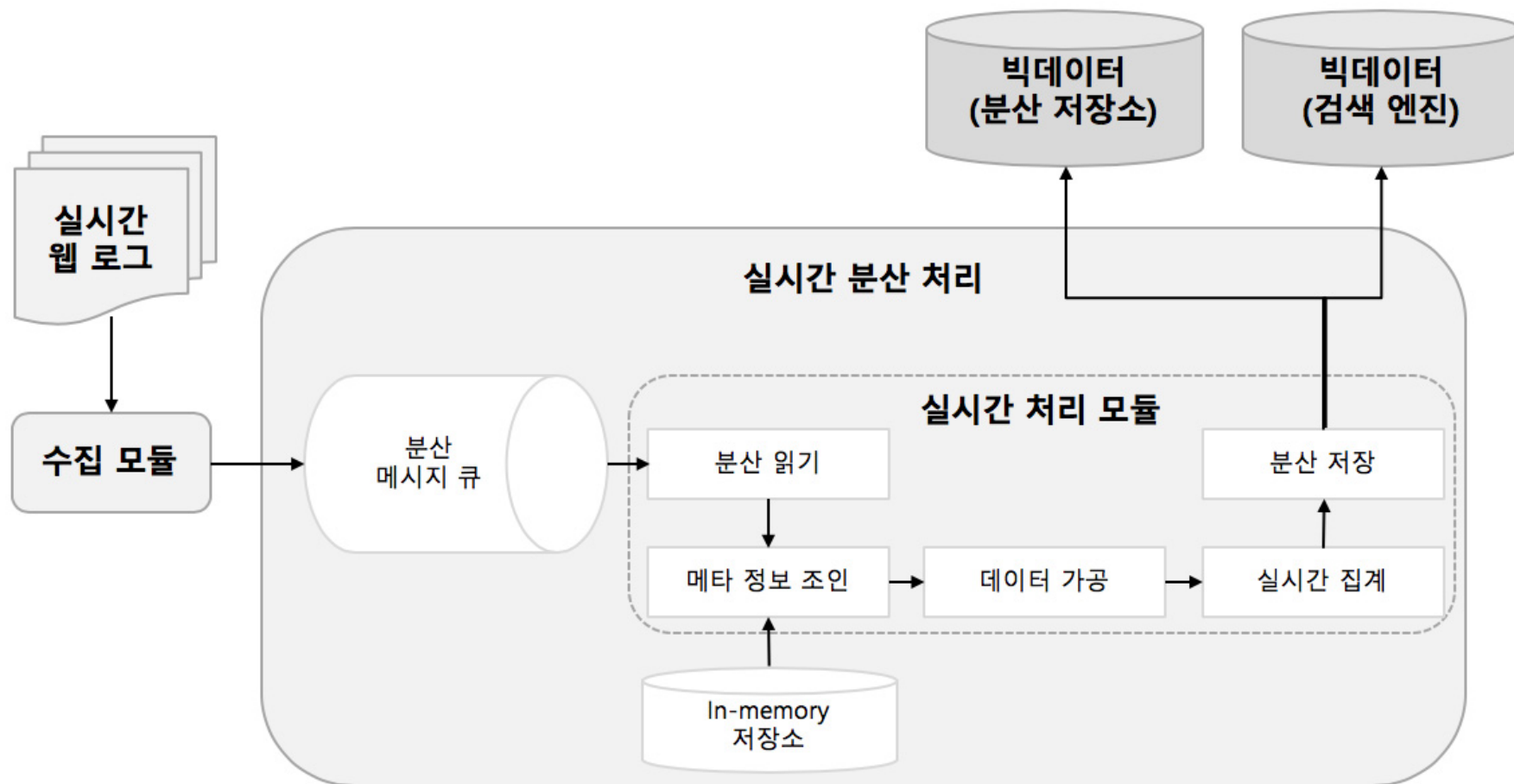


01. 데이터 분석 프로젝트



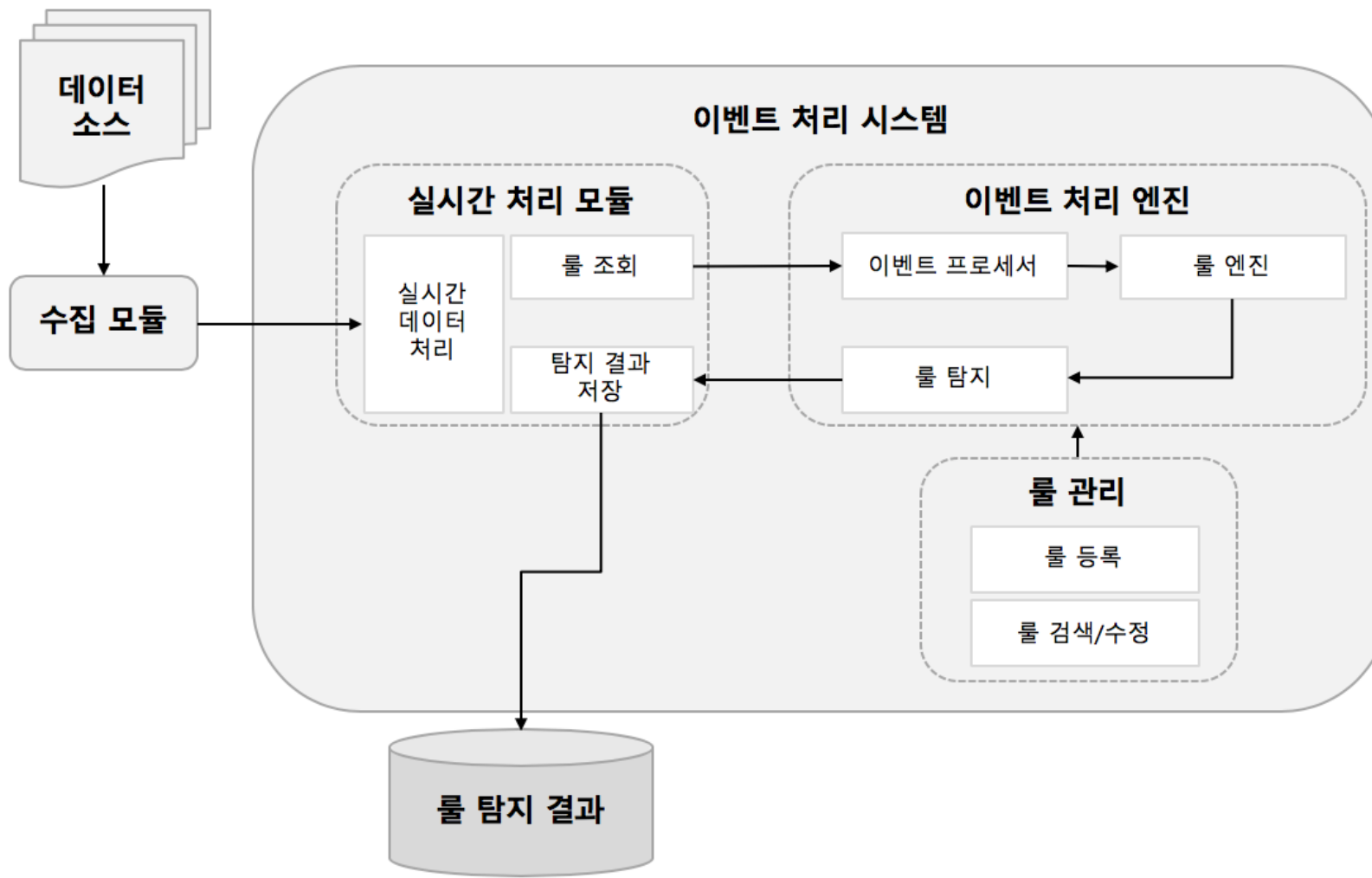


01. 데이터 분석 프로젝트





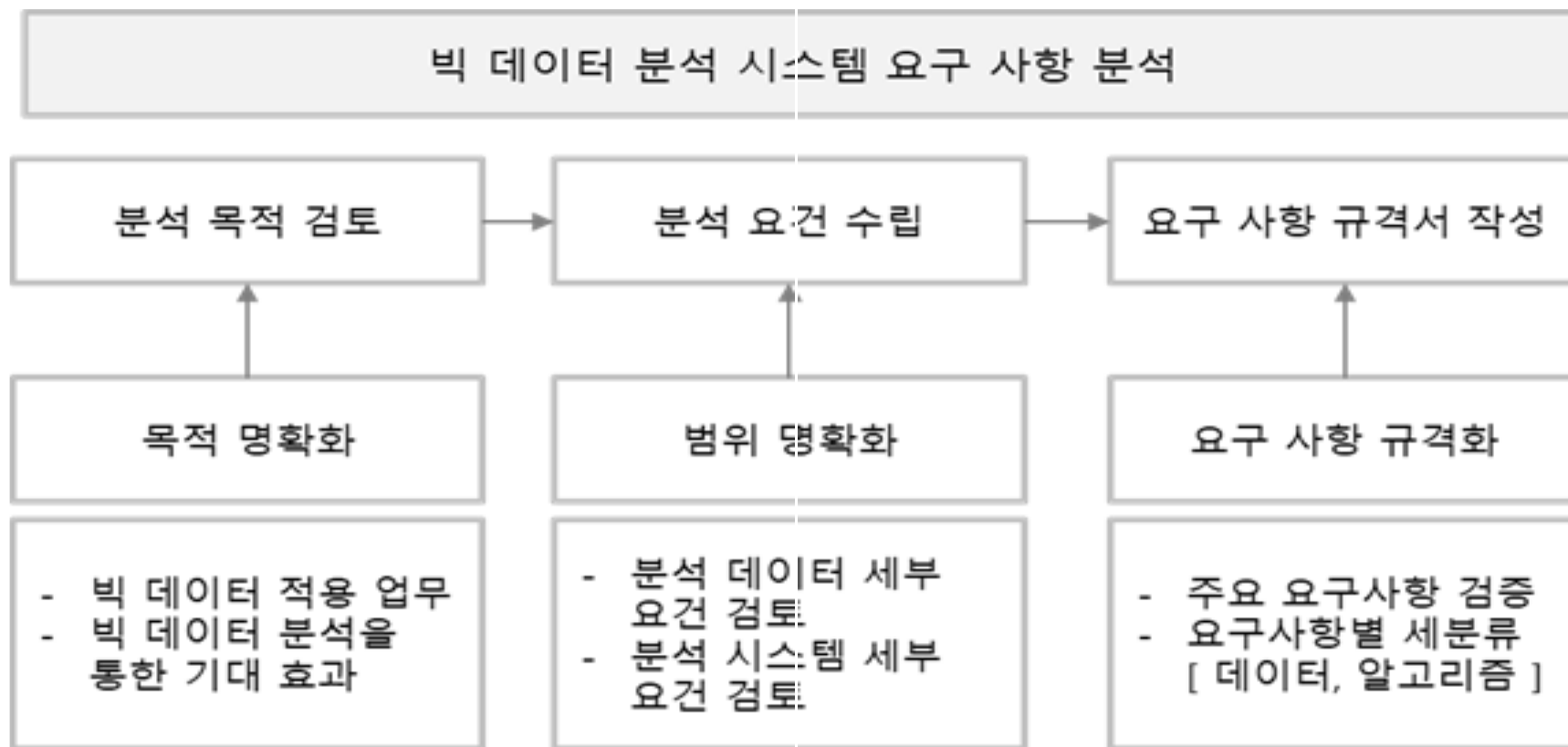
01. 데이터 분석 프로젝트





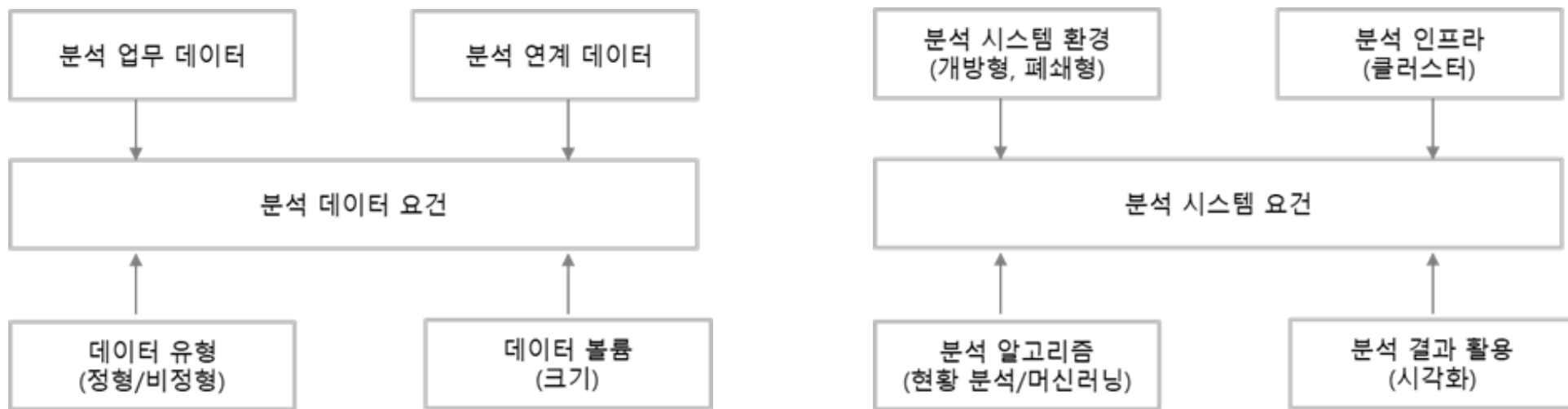
01. 데이터 분석 프로젝트

■ 빅데이터 분석 시스템 설계



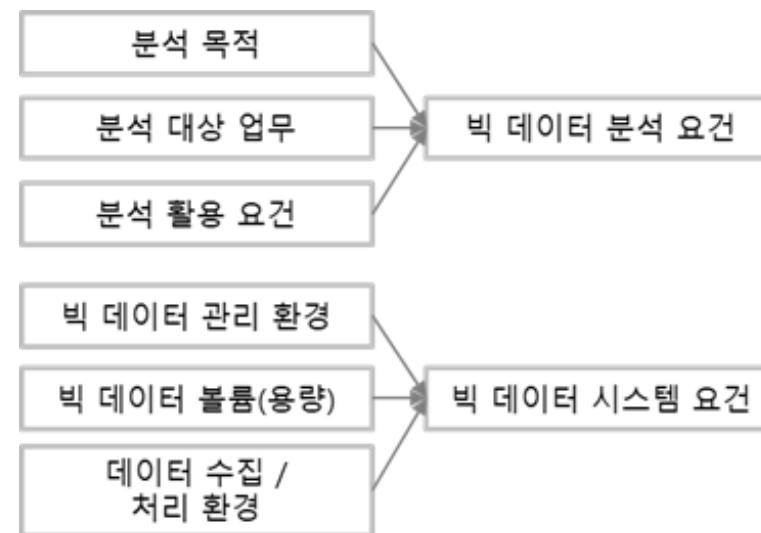
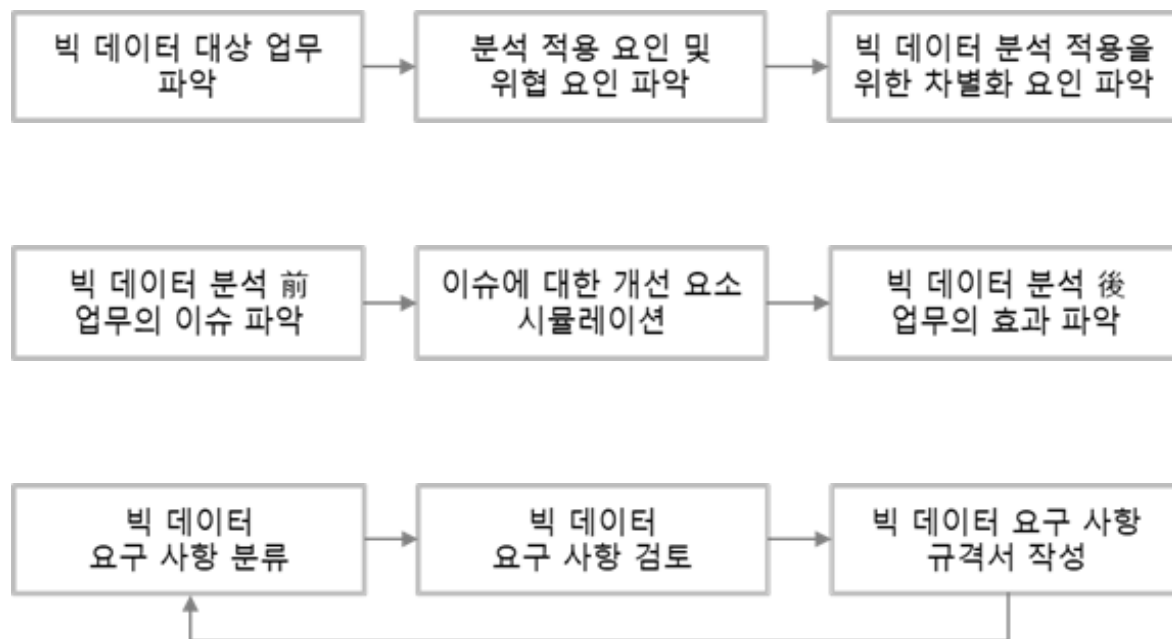


01. 데이터 분석 프로젝트





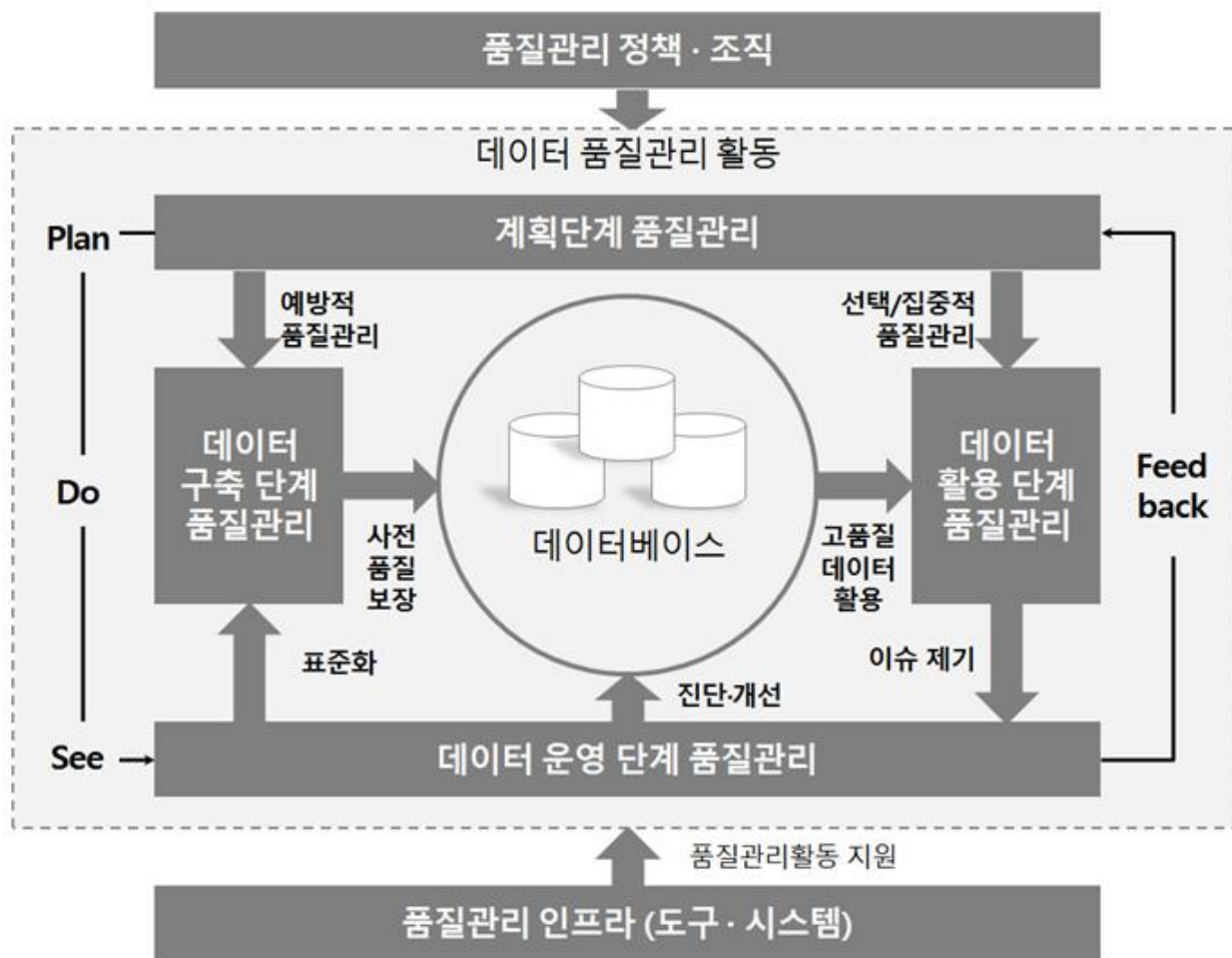
01. 데이터 분석 프로젝트





01. 데이터 분석 프로젝트

■ 빅데이터 품질관리 시스템 설계



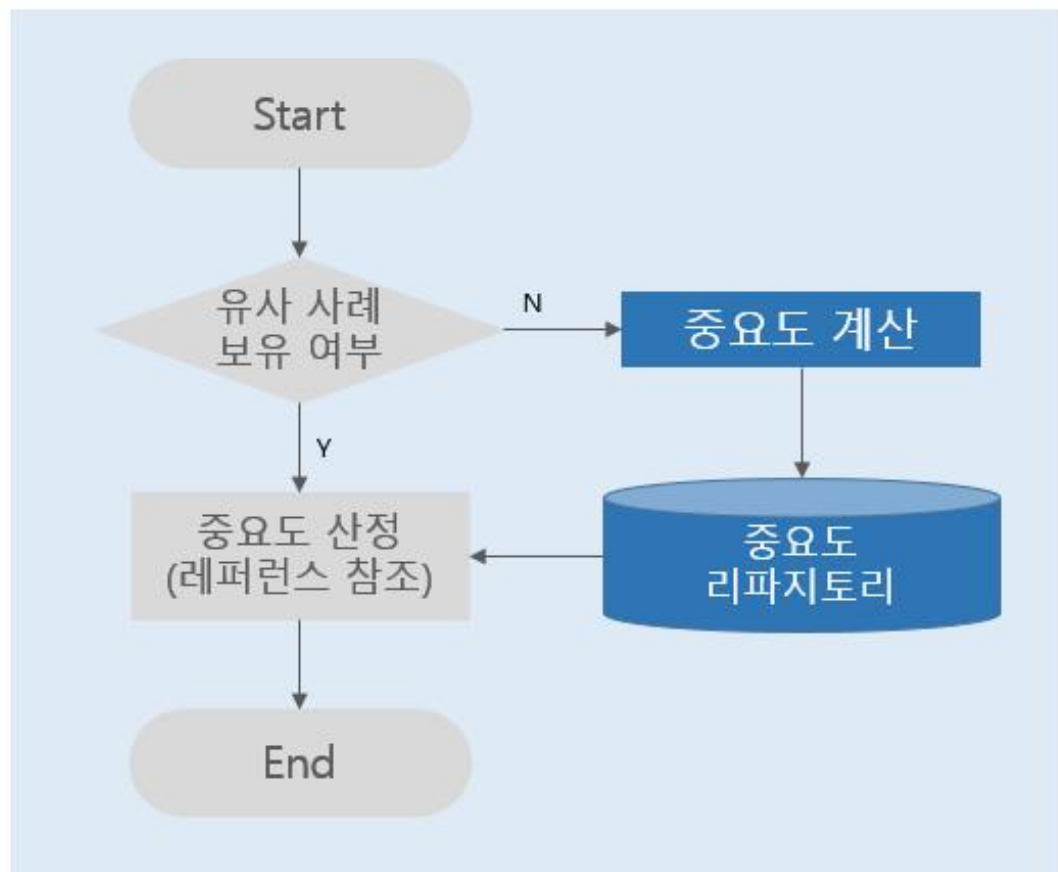


01. 데이터 분석 프로젝트

구성 요소		설명
단계별 데이터 품질 관리 활동	품질 관리 정책 및 조직	<ul style="list-style-type: none"> • 사용자 오류를 허용하는 데이터 수집 기준 마련 • 데이터 수집 예외 기준 마련
	구축 단계	<ul style="list-style-type: none"> • 데이터 분석 전 품질관리를 목표로 하며, 데이터 표준화, 데이터 구조의 일관성 확보 및 오류 데이터의 입력 방지 등의 활동으로 구성
	운영 단계	<ul style="list-style-type: none"> • 데이터의 품질진단 및 개선, 데이터 설계 및 표준 관련 문서 점검 • 데이터 설계 및 표준 문서 최신성 확보 및 이해관계자 공유 등의 활동으로 구성
	활용 단계	<ul style="list-style-type: none"> • 데이터 사용자들의 데이터 품질 오류 신고에 관한 사항을 접수하고 확인하는 활동으로 성과 평가를 통해 품질관리 목표를 재조정하는 활동으로 구성
품질 관리 도구 및 시스템		<ul style="list-style-type: none"> • 품질관리 활동을 어떻게 수행할 것인가에 대한 도구와 이를 지원하는 기술 및 시스템



01. 데이터 분석 프로젝트

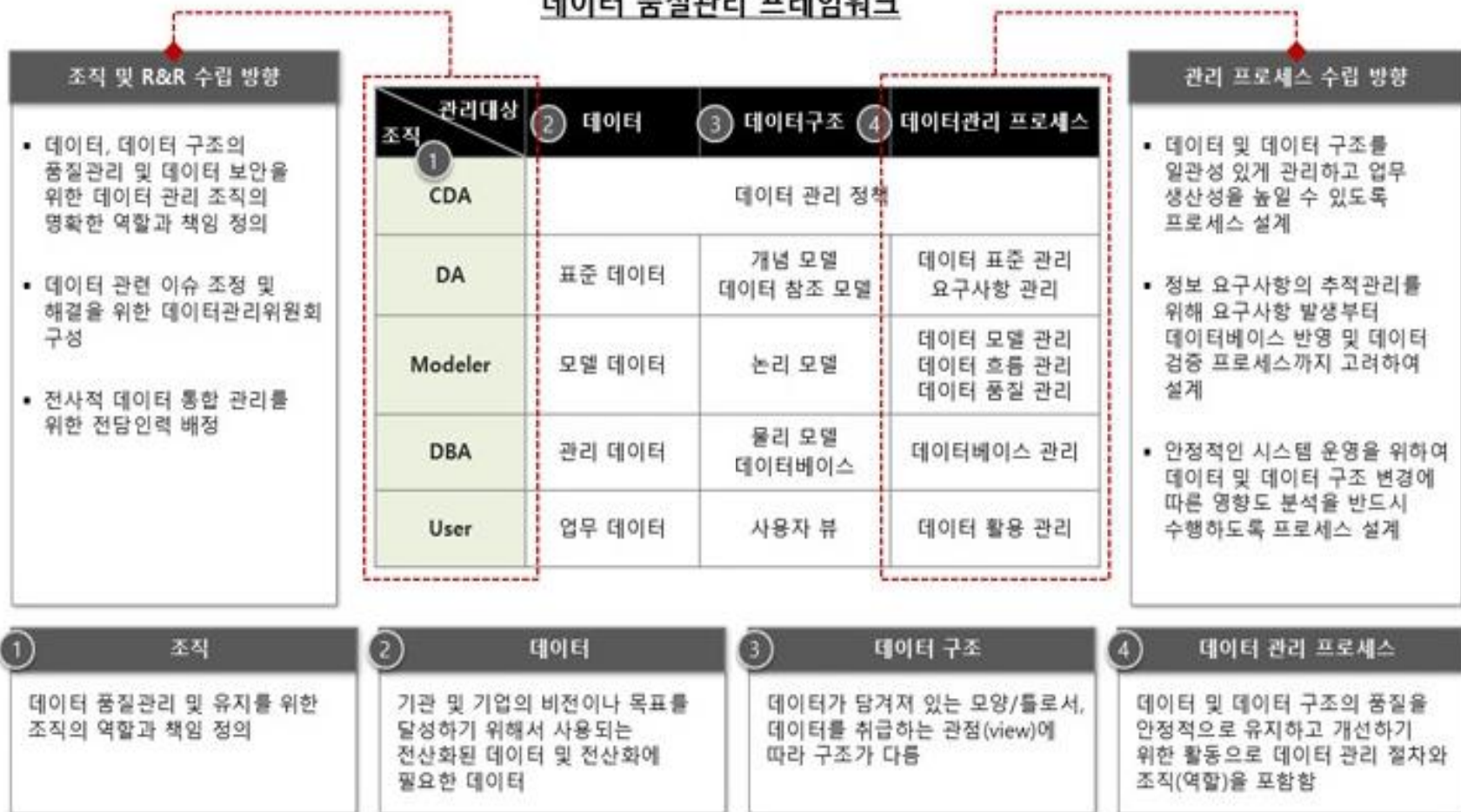


품질 기준	중요도
기능성	20%
신뢰성	10%
사용성	30%
효율성	20%
이식성	10%



01. 데이터 분석 프로젝트

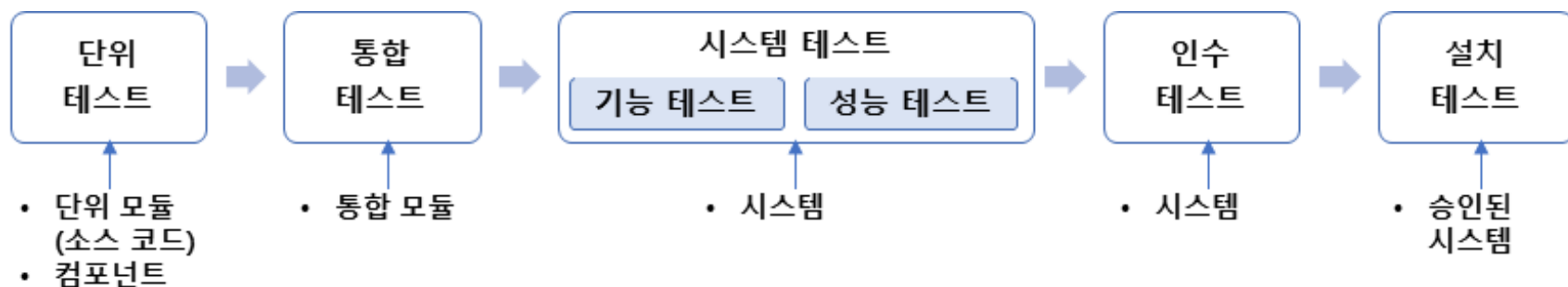
데이터 품질관리 프레임워크





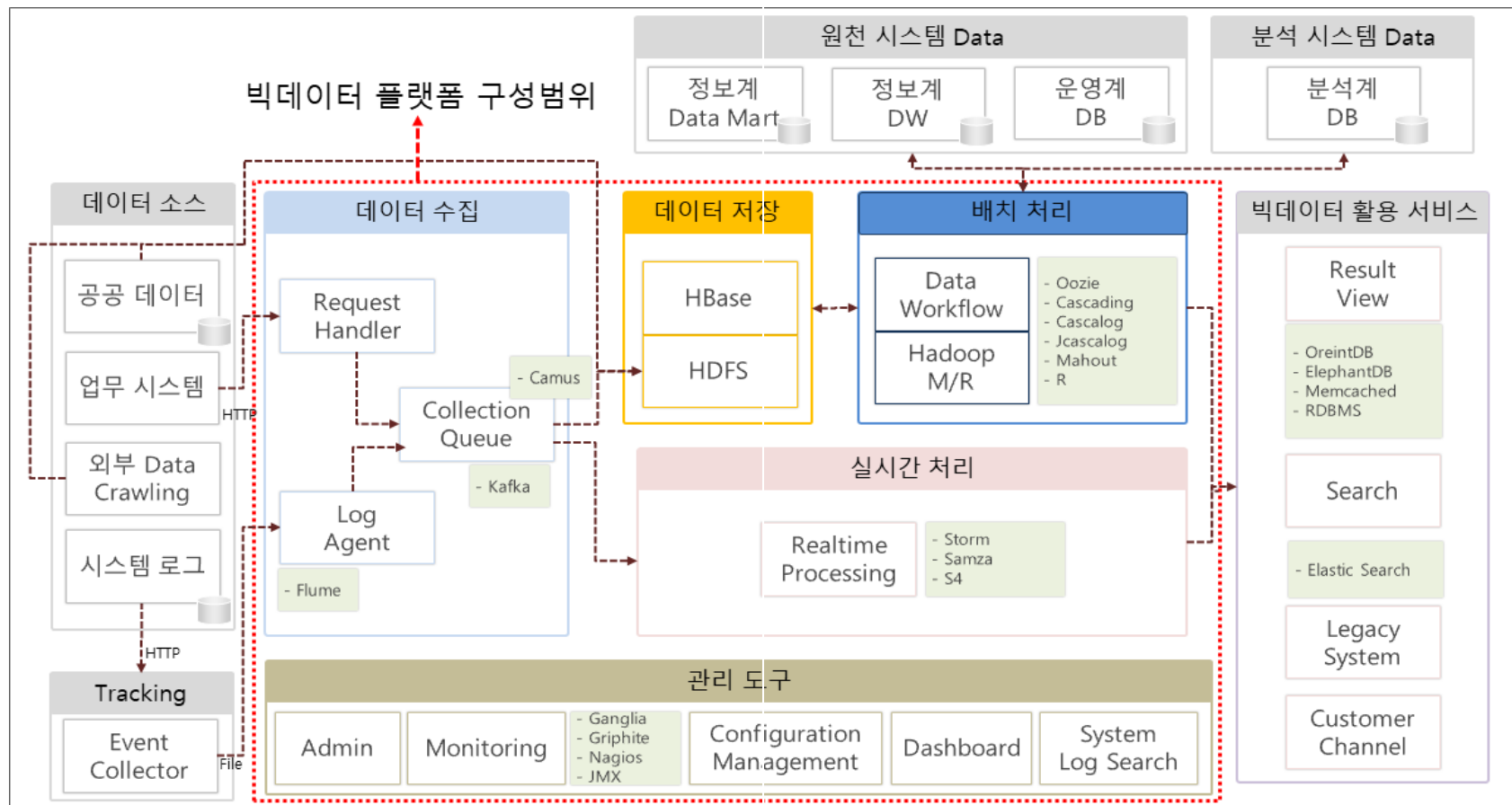
01. 데이터 분석 프로젝트

■ 빅데이터 플랫폼 테스트 설계





01. 데이터 분석 프로젝트



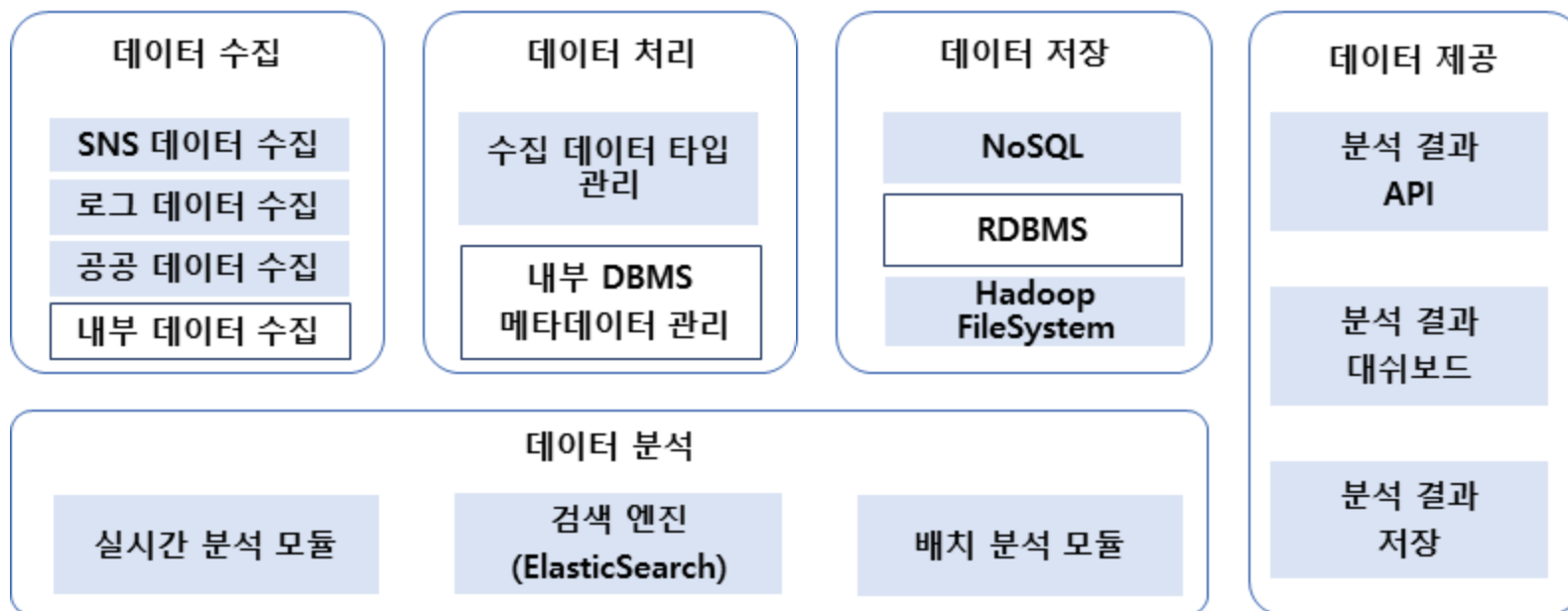


01. 데이터 분석 프로젝트

빅 데이터 플랫폼 테스트 범위

테스트 대상

테스트 제외





01. 데이터 분석 프로젝트

업무 프로세스 목록

업무 ID	업무명	담당자	작성일
001	회원가입	김민준	2023.01.01
002	로그인	김민준	2023.01.01
003	회원탈퇴	김민준	2023.01.01
004	비밀번호찾기	김민준	2023.01.01
005	회원정보수정	김민준	2023.01.01

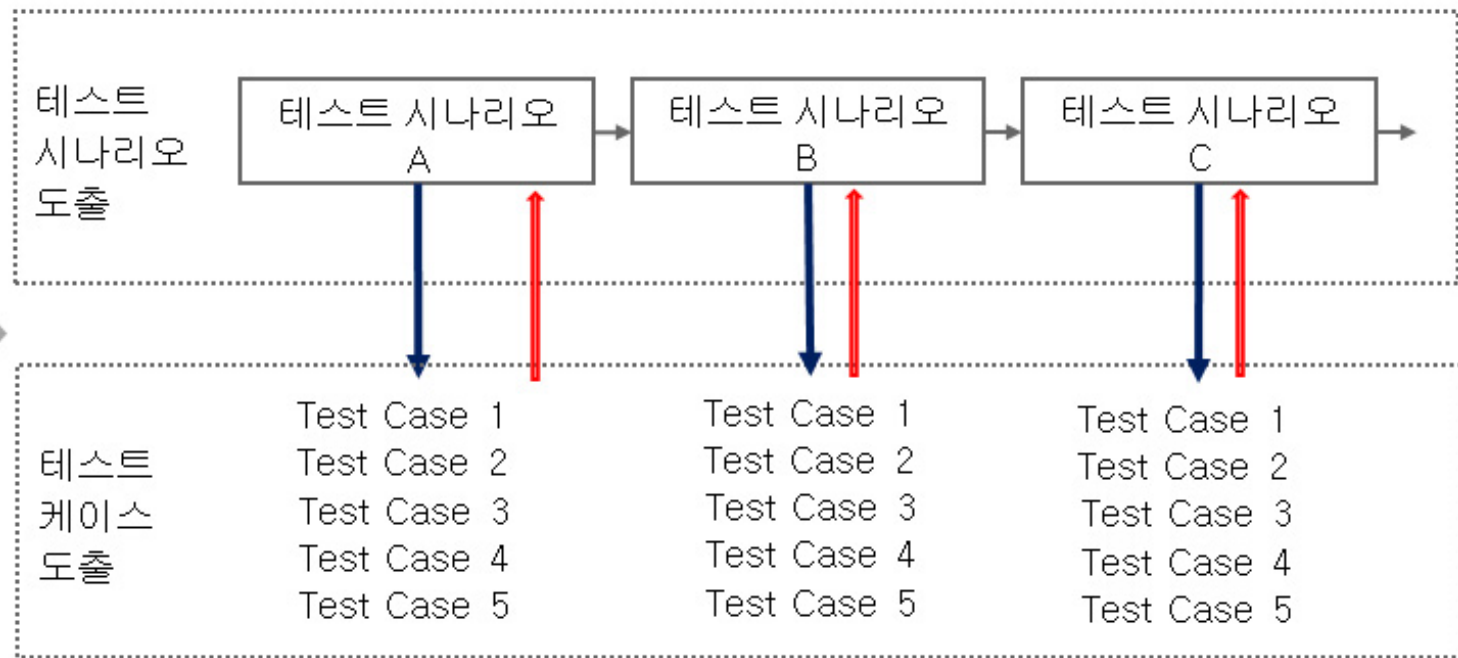


요구사항 정의서

요구사항 ID	요구사항명	우선순위	담당자	작성일
001	회원가입	고	김민준	2023.01.01
002	로그인	고	김민준	2023.01.01
003	회원탈퇴	중	김민준	2023.01.01
004	비밀번호찾기	중	김민준	2023.01.01
005	회원정보수정	중	김민준	2023.01.01

메뉴 구성도

메뉴 ID	메뉴명	상위메뉴	작성일
001	회원가입	회원관리	2023.01.01
002	로그인	회원관리	2023.01.01
003	회원탈퇴	회원관리	2023.01.01
004	비밀번호찾기	회원관리	2023.01.01
005	회원정보수정	회원관리	2023.01.01





01. 데이터 분석 프로젝트

■ 빅데이터 기술 플랫폼 운영 정책



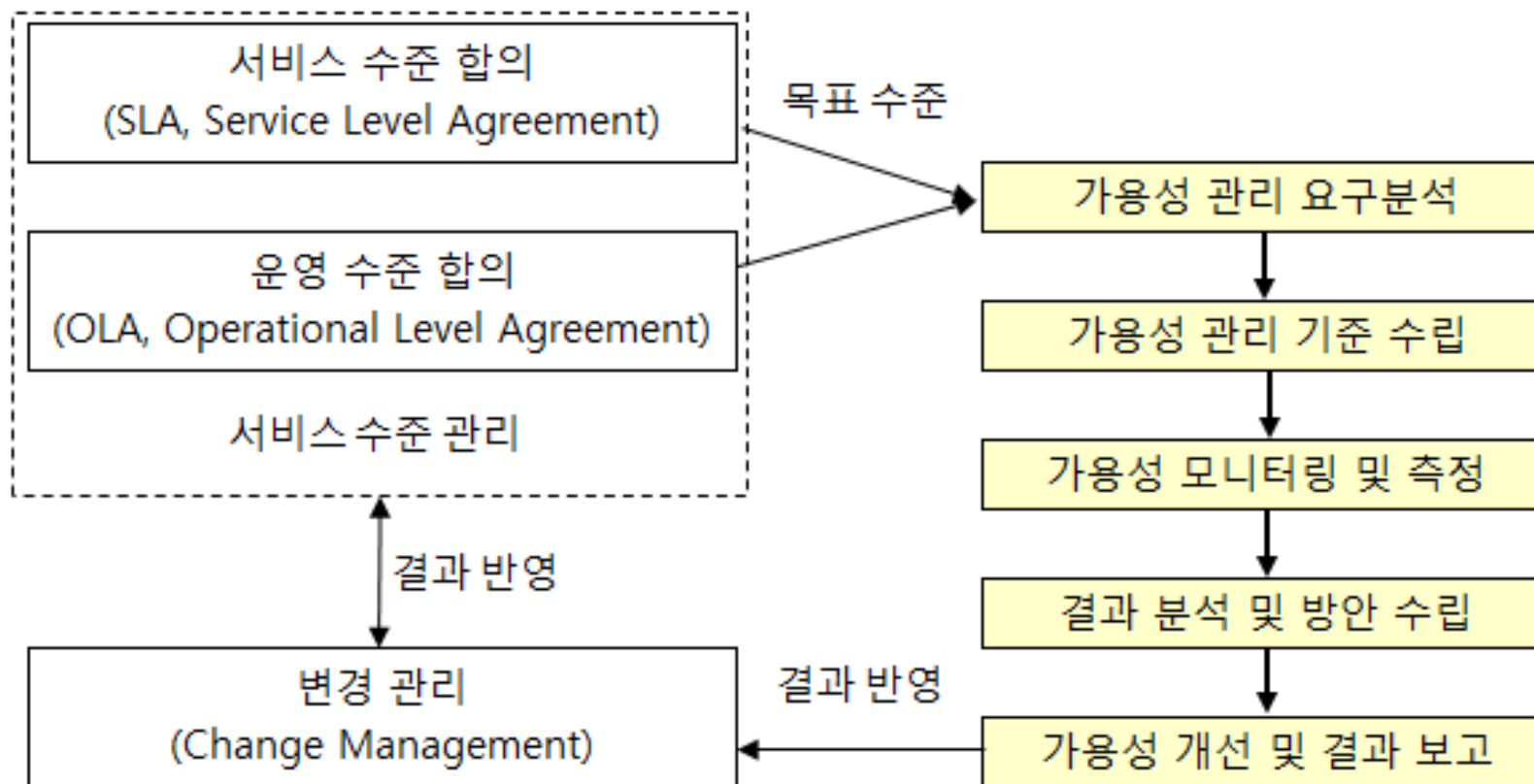


01. 데이터 분석 프로젝트



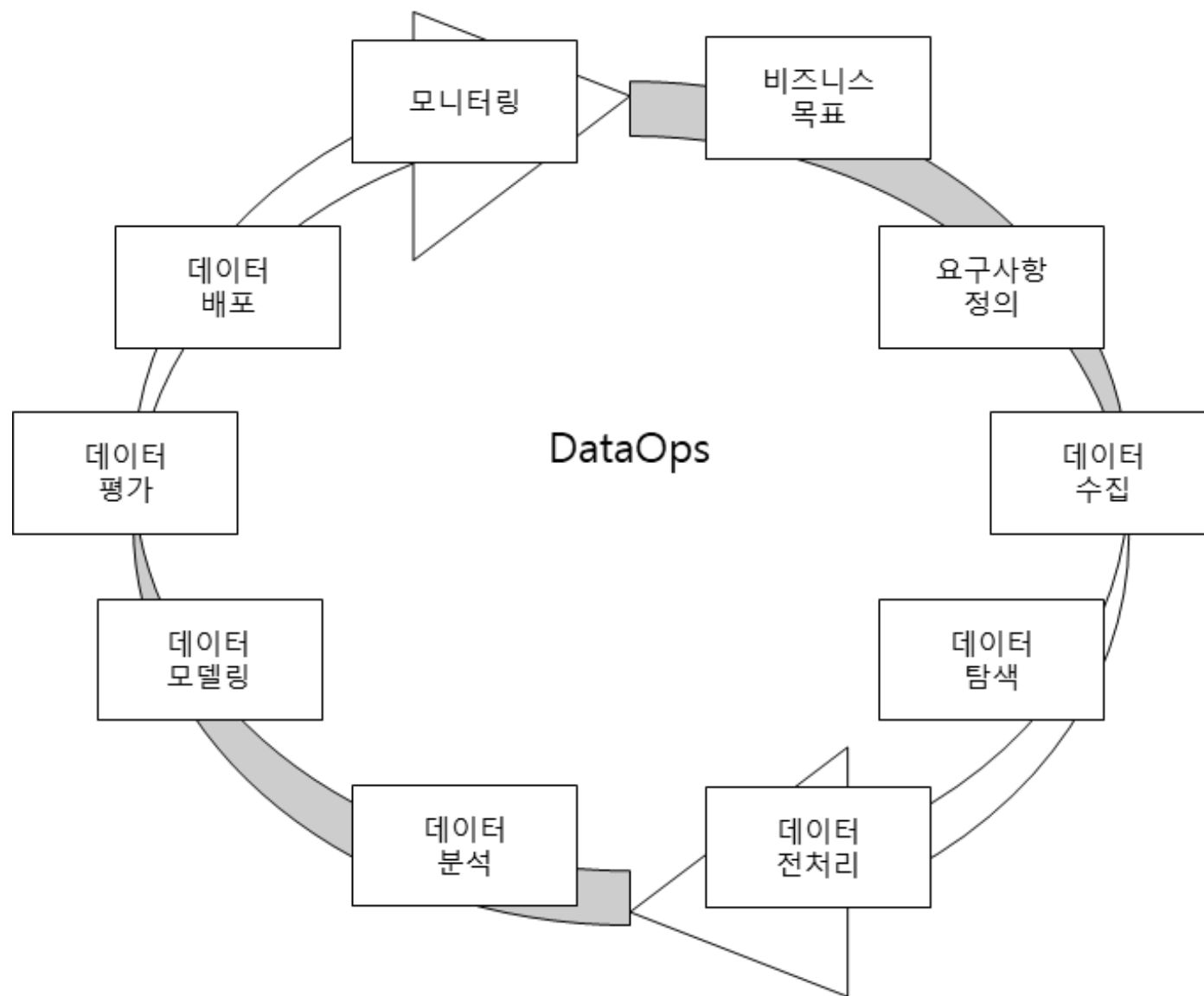


01. 데이터 분석 프로젝트



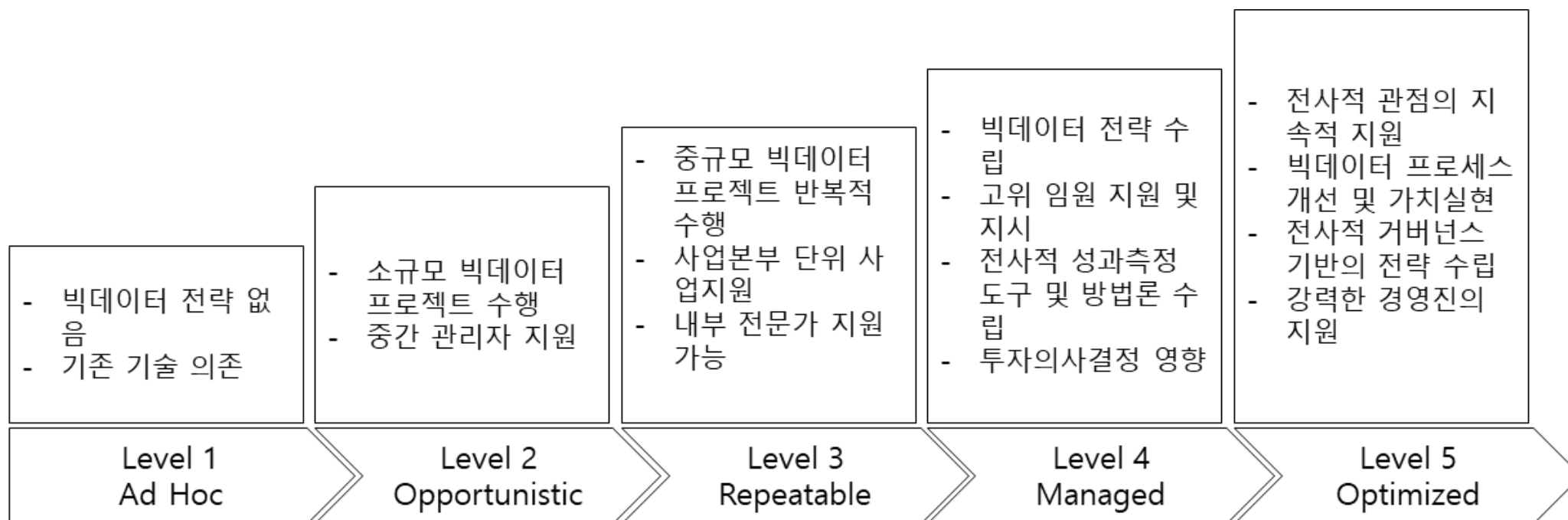


01. 데이터 분석 프로젝트



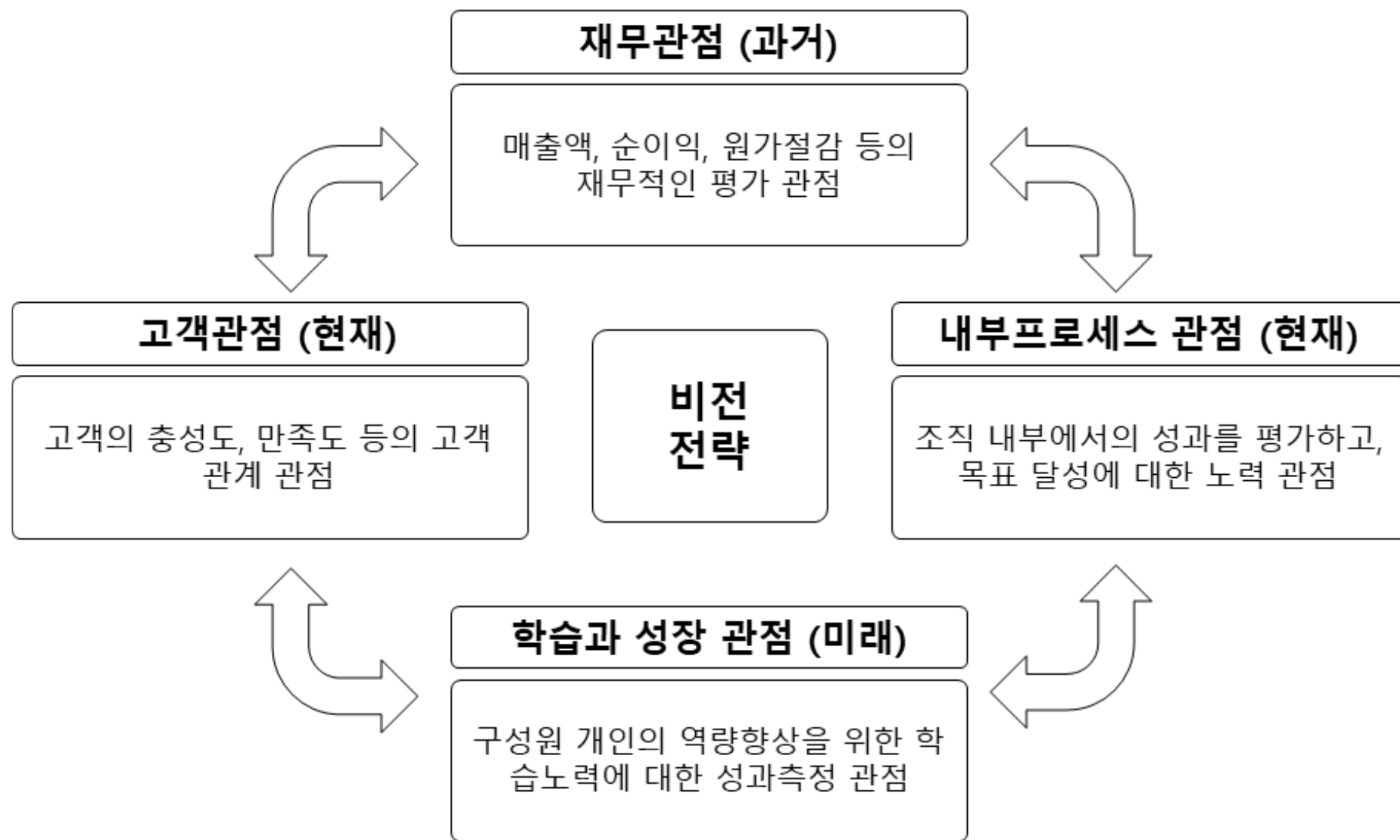


01. 데이터 분석 프로젝트





01. 데이터 분석 프로젝트





01. 데이터 분석 프로젝트

구분	프로세스	산출물
상위 조직	<pre> graph LR 1.1[1.1 상위 전략목표] --> 1.2[1.2 핵심성공요인 (CSF)] 1.2 --> 1.3[1.3 주요성과지표 (KPI)] 1.1 -.-> 2.1[2.1 상위 전략목표] 1.2 -.-> 2.2[2.2 핵심성공요인 (CSF)] 1.3 -.-> 2.3[2.3 주요성과지표 (KPI)] </pre>	<ol style="list-style-type: none"> 1. 상위 조직 미션 및 비전 2. 상위 조직 핵심전략 3. 핵심 전략별 CSF 4. 상위 조직 전략맵 5. 상위 조직 KPI
중간 조직	<pre> graph LR 2.1[2.1 상위 전략목표] --> 2.2[2.2 핵심성공요인 (CSF)] 2.2 --> 2.3[2.3 주요성과지표 (KPI)] 2.1 -.-> 3.1[3.1 상위 전략목표] 2.2 -.-> 3.2[3.2 핵심성공요인 (CSF)] 2.3 -.-> 3.3[3.3 주요성과지표 (KPI)] </pre>	<ol style="list-style-type: none"> 1. 중간조직 전략 2. 중간조직 전략별 CSF 3. 중간조직 전략맵 4. 중간조직 KPI
하부 조직	<pre> graph LR 3.1[3.1 상위 전략목표] --> 3.2[3.2 핵심성공요인 (CSF)] 3.2 --> 3.3[3.3 주요성과지표 (KPI)] </pre>	<ol style="list-style-type: none"> 1. 하부조직 CSF 도출 2. 하부조직 KPI 도출

- 핵심성공요인 (CSF : Critical Success Factor)
- 주요성과지표 (KPI : Key Performance Indicator)

Any Questions?

