**GS리테일**

**“데이터 거버넌스 체계 수립 및 품질진단 ~”**

**프로젝트 제안요청서**

**2022.06.XX**

목차

[1 프로젝트 개요 3](#_Toc106378949)

[1.1 프로젝트명 및 추진일정 3](#_Toc106378950)

[1.2 추진 배경 3](#_Toc106378951)

[1.3 목적 3](#_Toc106378952)

[2 추진 방향 및 전략 4](#_Toc106378953)

[2.1 추진 방향 4](#_Toc106378954)

[2.2 기대 효과 5](#_Toc106378955)

[2.3 추진전략 5](#_Toc106378956)

[3 업무 제안 범위 및 결과물 6](#_Toc106378957)

[3.1 컨설팅 요구사항 6](#_Toc106378958)

[3.2 기능 요구사항 8](#_Toc106378959)

# 프로젝트 개요

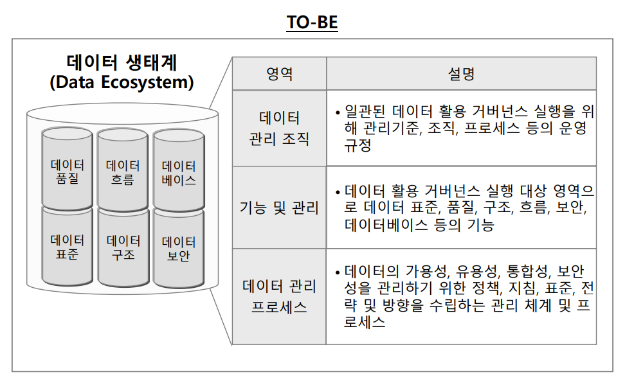
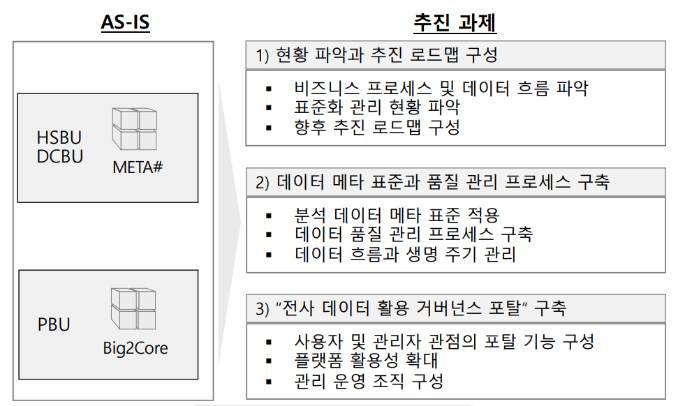
## 프로젝트명 및 추진일정

* 프로젝트명:
* 프로젝트 기간: 2022.07 ~ 2022.12 (6개월)

## 추진 배경

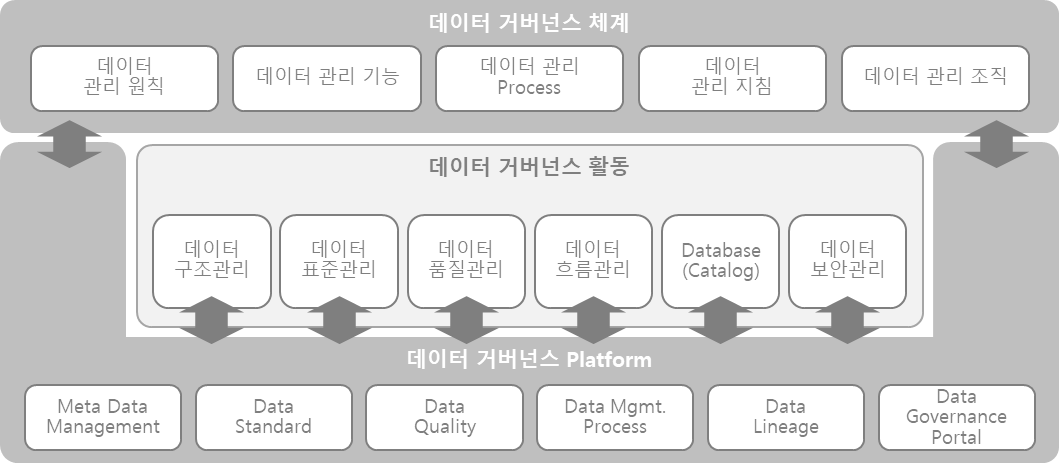
* 현재 유수의 기업들이 DT, DX 전략을 통한 데이터 기반의 의사결정 플랫폼을 구축하고 있으나, 당사는 다양한 사업 수행과 합병 등으로 방대한 데이터를 보유하고 있음에도 불구하고 유의미한 데이터를 도출하고 활용하기 위한 전사 차원의 데이터 거버넌스 (정책, 프로세스, 조직) 체계가 미흡하여 통합된 데이터 환경에서의 다양한 활용을 위해 많은 노력과 비용이 발생함
* 따라서, 동일 데이터 속성에 일관된 의미를 가진 데이터 품질을 확보한 상태로 지속적으로 유지하고, 데이터 분석을 통한 인사이트 도출과 미래 전략 수립을 뒷받침하기 위한 전사 데이터 거버넌스 플랫폼 도입이 요구

## 목적

* 정보 분석에 필요한 표준화된 데이터 구성 정보를 손쉽게 파악하고, 조회하여 정보 분석 데이터 구성 정보를 신속하게 획득하고 이해하여 활용할 수 있도록 “전사 데이터 거버넌스 플랫폼”을 구축, GS Retail 데이터 생태계(Data Ecosystem)을 구성하는 것이 목적임
* 이를 위하여 1) 현행 각 사업별 “비즈니스 데이터 흐름과 표준화 관리 현황 파악 및 추진 로드맵”을 구성하고 이를 기반으로 2) 정보 분석 관점의 “데이터 메타 표준과 품질 관리 프로세스 구축”하여 데이터 구조와 라이프사이클에 대한 관리 정책과 프로세스를 정립함. 이를 통하여 전사 임직원들이 손쉽게 접근, 활용하고, 거버넌스 관리 운영 조직이 데이터 변화 관리에 신속하고 능동적으로 대처할 수 있도록 3) “전사 데이터 활용 거버넌스 포털”을 구축하고자 함

# 추진 방향 및 전략

## 추진 방향

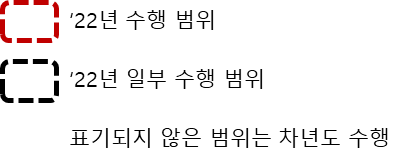
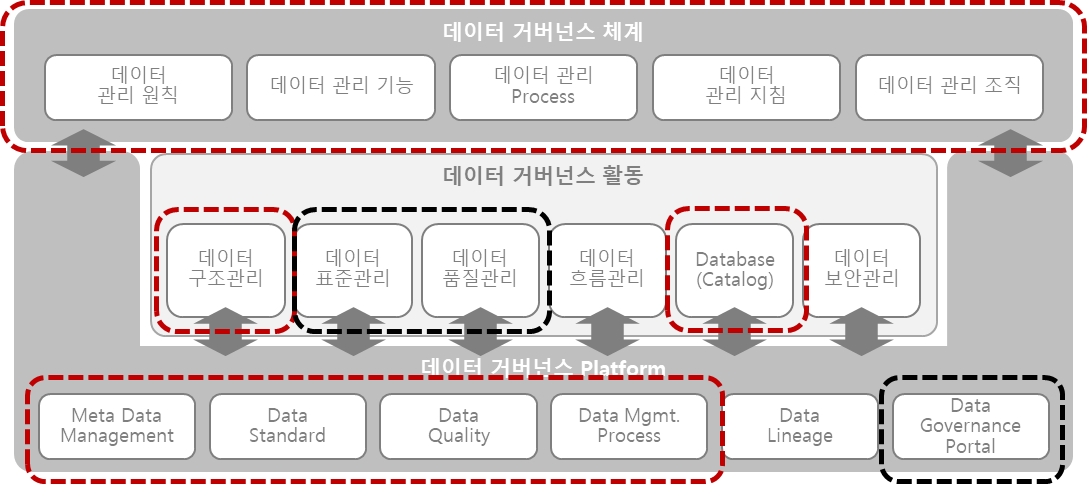


* 당사 각 사업부 및 정보분석 시스템에서 발생되고 있는 데이터의 비즈니스적 흐름을 이해하고, 데이터 속성 표준 지침을 수집, 관리 수준을 파악하여 “전사 데이터 거버넌스 플랫폼” 구성 범위를 정의하고 상세 추진 과제와 일정 계획을 확정함
* 현행 데이터 표준 지침 자료를 기반으로 데이터 분석과 활용 시 일관성 있는 데이터 구조와 의미를 제공할 수 있는 데이터 활용 거버넌스 메타 표준 구성 방향성을 수립함
* 데이터 수집과 활용에 대한 보안 체계를 검토하고 (개인정보 보유 범위 등), 조직별 활용 데이터 접근 권한 체계 지침에 따라 구성함
* 데이터 구성 변화 관리 프로세스를 수립하고 처리 방안에 대한 유관 부서와의 협의하여 정립함
* 분석 데이터 메타 표준 시스템 도입을 추진하고, 정보 분석 관점의 단어, 용어, 도메인에 대한 표준화와 더불어 데이터 구조와 속성의 의미를 명확하게 정의하여, 데이터 카탈로그(Data Catalog)를 구성함
* 각 사업부별로 관리되고 있는 데이터 속성 표준 코드를 통합하고 정제하여 정보 분석 활용에 적용할 수 있도록 코드 표준화와 원천(소스) 데이터 매핑 규칙을 시스템에 적용함
* 데이터 품질 관리를 위하여, 수집하고 활용하는 데이터들의 메타 표준 준수 여부를 파악할 수 있도록 주기적인 품질 관리 자동화 수행과 모니터링 체계를 설계하고 품질 진단 파일럿을 실행한 진단 결과에 대한 개선 사항을 유관 부서에 공유하고 대응 조치 협업 프로세스를 구성하여, 데이터의 품질을 지속적으로 확보할 체계를 마련함
* 원천(소스)부터 활용까지 데이터의 흐름을 정의하고, 이러한 흐름 정보를 시스템 내 구성하여 데이터 생명주기에 대한 모니터링 체계를 설계함

## 기대 효과

* 데이터 활용 관점에서 분석 지표 의미의 명확성 제공과 더불어, 검증되고 일관된 데이터 품질을 확보하여 조직 구성원들의 데이터에 대한 이해도와 활용 능력을 증대할 수 있음
* 활용 데이터의 신속한 업무 처리 프로세스를 적용하여, 구성원들의 전략적 사업 기획과 추진을 위한 소요시간 (Lead Time)을 단축시키고, DT, DX에 기반한 미래 사업 발굴과 수행 성공률을 높여 당사의 시장 경쟁력과 인지도를 강화할 수 있는 토대를 마련할 수 있음

## 추진전략



* 현재 GS리테일의 데이터 거버넌스는 원천 시스템을 포함하여 전사에 걸쳐 적용하지 않고, 분석 관점에서 데이터의 활용 관점에서 진행되므로, 내부의 변화관리까지 고려하여 단계적인 적용을 추진함
* 따라서, ‘22년은 GS리테일 데이터 거버넌스 도입의 첫 단계로, 1) 기본적인 Data 관리 기능(시스템/솔루션)의 구축, 2) 향후 데이터 거버넌스 활동의 추진 계획 수립, 그리고 3) 데이터 거버넌스의 핵심 활동인 데이터 표준화 및 데이터 품질진단의 시범적 수행을 추진하는 것을 목표로 사업을 추진함
  + 기본적인 Data 관리 기능의 구축을 위하여, 시급성과 데이터 거버넌스 Platform의 기능 우선순위를 고려하여 단계별 데이터 거버넌스 시스템 도입
  + 시스템 도입 일정 및 거버넌스 환경 성숙도에 따라 향후 추진되어야 하는 데이터 거버넌스 도입 로드맵 수립
  + 전체 데이터 거버넌스 Platform의 기능을 고려한 데이터 거버넌스 포털 요구사항 정의 및 개념 설계
  + 데이터 활용도를 고려하여 핵심 분석 시스템에 대하여 데이터 표준화 및 표준관리 적용
  + 핵심 데이터를 대상으로 한 Pilot 데이터 품질 진단 수행

# 업무 제안 범위 및 결과물

## 컨설팅 요구사항

|  |  |
| --- | --- |
| 요구사항 구분 | 현황 분석 |
| 요구사항 상세구분 | 비즈니스 프로세스 및 데이터 흐름 파악 |
| 세부내용 | * 현행 데이터 분석 * 구체적인 진단 기법, 방법론, 도구 등을 활용 * 데이터 흐름 파악 |
| 산출물 | * 현행 데이터 분석서 |

|  |  |
| --- | --- |
| 요구사항 구분 | 현황 분석 |
| 요구사항 상세구분 | 표준화 관리 현황 파악 |
| 세부내용 | * 현행 데이터 표준 가이드, 환경분석 * 현행 데이터 표준 현황 분석 * 표준/비표준DB 파악 * 정보계 통합 관점에서 표준 단어, 용어, 도메인, 코드 분석 * 데이터 표준화 요구사항 수집, 정의 * 모델과 표준의 연계 요구사항 * 메타시스템에서 관리할 관리항목 도출 및 시스템반영 요건 * 표준화 수행을 위한 요구사항 도출, 수집, 정의 |
| 산출물 | * 현행 데이터 분석서, 요구사항 정의서 |

|  |  |
| --- | --- |
| 요구사항 구분 | 로드맵 수립 |
| 요구사항 상세구분 | 향후 추진 로드맵 수립 |
| 세부내용 | * 시스템(포털, 메타, 품질) 구축 전략 및 계획수립 * 처리 프로세스 정의 * 구축 일정 수립 * Data Lake 기반 정보계 구성 표준 방향성 수립 * 데이터 거버넌스 추진 상세 계획 수립 * 데이터 표준화, 품질 확산 * 데이터 거버넌스 포털 커스터마이징 * Data Lineage 도입 |
| 산출물 | * 데이터 거버넌스 이행계획서 |

|  |  |
| --- | --- |
| 요구사항 구분 | 데이터 메타 표준과 품질관리 프로세스 구축 |
| 요구사항 상세구분 | 핵심 시스템 데이터 표준화 수행 |
| 세부내용 | * 표준 단어, 용어, 도메인, 코드 사전 제작 * 현행 메타시스템을 분석하여 정보계 통합 관점의 메타 표준 사전(단어, 도메인, 용어, 코드) 제작 * 표준 단어, 용어, 도메인, 코드 유효성 검증 * 현행 메타시스템과 매핑 규칙 정의 * 현행 표준 단어, 용어, 도메인, 코드와 매핑 규칙 정의 * 매핑 지속 관리 프로세스 수립 * 데이터 표준화 수행 * 도입된 메타 표준 시스템을 활용한 데이터 표준화 수행 |
| 산출물 | * 표준 정의서 |

|  |  |
| --- | --- |
| 요구사항 구분 | 데이터 메타 표준과 품질관리 프로세스 구축 |
| 요구사항 상세구분 | 핵심 데이터 품질진단 수행 |
| 세부내용 | * 데이터 진단 대상 범위 * 현행 데이터 품질 현황 및 이슈 수집 분석   + - * 시스템 별 데이터 표준 상이       * 시스템 별 상품코드, 상품분류 등 상이       * 동일 바코드에 상품, 상품분류 등 상이       * 상품의 필수 속성 부재 ( 상품분류 등 ) * 데이터 품질 요구사항 수집 및 정의 * 핵심 데이터 진단 대상 선정 * 데이터 진단 규칙 정의 * 데이터 품질관리 기준 정의 (도메인, 필수값, 중복값 등) * 데이터 업무규칙 수집 및 진단 규칙 정의 * 데이터 품질진단 수행 및 결과 도출 * 도입된 품질관리 시스템을 활용한 진단 수행 * 진단결과를 분석하여 문제점 제시, 시사점 도출 등 보고서 작성 * 데이터 품질 개선 조치 * 품질진단 결과 데이터 확인 * 데이터 품질 개선 지원 |
| 산출물 | * 진단대상 정의서, 품질진단 기준 정의서, 업무규칙 정의서, 품질진단 결과보고서 |

|  |  |
| --- | --- |
| 요구사항 구분 | 데이터 메타 표준과 품질관리 프로세스 구축 |
| 요구사항 상세구분 | 상품 표준화 지원 |
| 세부내용 | * 정보계 통합 상품 표준화 지원 * 통합 상품 분류 분석 지원 * 상품 분류 미정의 상품의 처리 방향 지원 |
| 산출물 |  |

|  |  |
| --- | --- |
| 요구사항 구분 | 데이터 거버넌스 포털 구축 |
| 요구사항 상세구분 | 데이터 거버넌스 포털 요구사항 정의 및 개념 설계 |
| 세부내용 | * 데이터 거버넌스 포털 요구사항 정의 * 데이터 거버넌스 솔루션 기능 분석 및 필요한 부가 기능 정의 * 데이터 거버넌스 포털 기능 요구사항 도출 * 데이터 거버넌스 포털 기능 화면 설계 * 기능 요구사항에 따른 커스터마이징 화면 설계 |
| 산출물 | * 요구사항정의서, 데이터 거버넌스 포털 개념 설계서 |

## 기능 요구사항

|  |  |
| --- | --- |
| 요구사항 구분 | 시스템 도입 및 기능 |
| 요구사항 상세구분 | 데이터 거버넌스 포털 |
| 세부내용 | * 지원기능 * 메타 표준 시스템, 품질관리 시스템과 연동 * 품질 및 메타 정보를 동일 시스템 관점으로 통합 대시보드 제공 * 로그인 사용자 기반 메타 및 품질 관리 영역의 To-Do 리스트 관리 * 시스템 별 메타 변경 사항, 품질 현황 추이 관리 * 통합 검색 기능 지원 * 시스템, 사용자 등 기반 정보에 대한 시스템 통합 연계 관리 * 사용자 및 권한 관리 |
| 산출물 | * 데이터 거버넌스 포털 설명서 |

|  |  |
| --- | --- |
| 요구사항 구분 | 시스템 도입 및 기능 |
| 요구사항 상세구분 | 메타 표준 시스템 |
| 세부내용 | * 데이터 표준관리 기능 * 데이터 표준관리 (등록, 수정, 삭제) * 단어, 용어, 도메인, 코드 등 표준사전 관리 * 표준 용어와 비표준 용어 매핑 관리 * 데이터 표준 정보 검색, 조회 * 실제 DB와 데이터 표준간 불일치(Gap) 관리 * 지원기능 * 워크플로우 관리 : 메타데이터의 등록, 변경 프로세스에 대한 워크플로우 정의 및 통제 * 사용자 관리 : 메타데이터에 대한 사용자 및 권한 등록 * 스케줄 관리 : 메타데이터 추출/적재, 메타데이터 검증 등의 작업에 대한 스케줄 관리 * 메타 변경 내역 조회 : 메타데이터 과거 변경 내역 조회 * 사용 통계 조회 : 메타데이터 사용 현황 조회 및 통계 분석 |
| 산출물 | * 메타 표준 시스템 설명서 |

|  |  |
| --- | --- |
| 요구사항 구분 | 시스템 도입 및 기능 |
| 요구사항 상세구분 | 품질관리 시스템 |
| 세부내용 | * 데이터 품질진단 대상 및 검증기준 관리 * 데이터 품질지표(완전성, 유효성, 정합성, 일관성 등 DQI) 관리 * 핵심항목(CTQ) 관리 * 품질검증 대상 추출 및 관리 기능 * 컬럼 분석 및 데이터 패턴 분석 기능 * 중복검증, 날짜검증, 코드검증, 관계검증 기능 * 업무규칙(BR) 관리 기능 * 데이터 품질검증 작업 관리 * 데이터 품질검증 수행 * 진행상황에 대한 모니터링 * 데이터 검증 스케줄 설계 및 실행 등 프로세스 관리 * 지원기능 * 워크플로우 관리 : 품질관리 분석, 진단, 통제 프로세스에 대한 워크플로우 정의 및 통제 * 사용자 관리 : 품질담당자에 대한 사용자 및 권한 등록 |
| 산출물 | * 품질관리 시스템 설명서 |