과제명 : 데이터거버넌스

데이터플랫폼본부 데이터거버넌스팀(’22.06.23)

|  |
| --- |
| < 목차 >  Ⅰ. 당사 현황  Ⅱ. 추진 방향  Ⅲ. 구체 추진 계획: 구체 과제, 일정, 인원, 비용 등  **Ⅳ. 최종 결과** |

1. 당사 현황

* **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 활용 거버넌스(정책, 프로세스, 조직) 체계가 미흡하여 통합된 데이터 환경에서의 분석과 활용을 위한 많은 노력과 비용이 발생하고 있음**

|  |  |  |
| --- | --- | --- |
| 구 분 | 내 용 | 구성 요소 |
| 데이터 운영 거버넌스 | * 시스템별로 산재해 있는 데이터 정보 요소에 대한 명칭, 정의, 형식, 규칙에 대한 원칙을 수립하여 이를 전사적으로 적용하는 것을 의미하는 비즈니스 처리를 위한 데이터 표준 관리 체계 * 비즈니스에 의해 발생하는 데이터의 표준 단어/용어/도메인을 정의하고 속성의 의미를 유일하게 확정하며, RDB 관계에 의한 데이터의 무결성, 정합성을 유지하기 위한 정책과 프로세스 * 트랜잭션 위주의 RDB 관계(OLTP) 모델에 기반한 정형 데이터 표준 및 품질 관리에 주로 활용함 | 비즈니스 운영 관점  데이터 메타  데이터 품질(DQ)  데이터 모델 |
| 데이터 활용 거버넌스 | * 데이터가 조직에 높은 가치를 제공할 수 있도록 데이터의 효과적인 획득, 관리 및 활용과 관련된 관행, 정책 및 역할의 집합을 의미함 * 데이터의 수집, 통합, 활용에 관련된 전략과 정책을 수립하고 이를 관리하기 위한 운영 조직을 구성하여 데이터를 통한 업무 처리 프로세스를 확립하는 체계를 의미하며, 이를 통하여 기업은 데이터의 관리 수준과 품질을 향상시켜 전략적 의사결정을 신속하게 할 수 있는 체계를 구성할 수 있음 * 정형/비정형 데이터의 활용(OLAP) 위주의 모델에 기반한 빅데이터 표준 및 수명 주기, 정책 관리에 주로 사용함 | 데이터 자산  활용 관점  데이터 메타  데이터 품질(DQ)  데이터 프로파일링  데이터 강화(결합)  데이터 리니지  데이터 활용 보안 |

[표 1. 데이터 거버넌스 구분]

* + 당사 내 분산 관리되고 있는 데이터 구성 환경으로 인하여 전사 차원의 전략적 의사결정에 필요한 유의미한 데이터 파악을 위해 과도한 소요시간(Lead Time)이 발생하고 있으며, 고객에게 데이터 분석 기반의 다양한 서비스를 발굴하여 신속하게 제공하기 어려움
  + 또한, 급변하는 시장 환경 변화에 따라 데이터는 지속적으로 변화하고 있으며 생성, 변환, 활용, 소멸에 이르는 데이터 라이프 사이클 관리 부족 등으로 인하여 일관성 있는 분석 데이터 품질을 지속적으로 유지하기 어려움
* **따라서, 동일 데이터 속성에 일관된 의미를 가진 데이터 품질을 확보한 상태를 지속적으로 유지하고, 데이터 분석을 통한 인사이트 도출과 미래 전략 수립을 뒷받침 하기 위한 전사 데이터 활용 거버넌스 플랫폼 도입이 시급한 과제로 판단 됨**
* **이를 위하여, 데이터 속성의 명확한 의미를 정의하고, 데이터의 발생부터 소멸까지의 흐름 파악과 지속적인 품질을 유지하면서, 임직원들의 데이터 이해 증진을 도모하기 위한 데이터 활용 거버넌스 운영 조직 구성이 필요 함**

1. 추진 방향

* **본 전략 과제는, 정보 분석에 필요한 표준화된 데이터 구성 정보를 손쉽게 파악하고, 조회하여 분석 데이터 구성 정보를 신속하게 획득하고 이해하여 활용할 수 있도록 “전사 데이터 활용 거버넌스 플랫폼”을 구축, GS리테일 전사 데이터 생태계(Data Ecosystem)을 구성하는 것이 목적 임**
* **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 포탈”을 구축하고자 함**



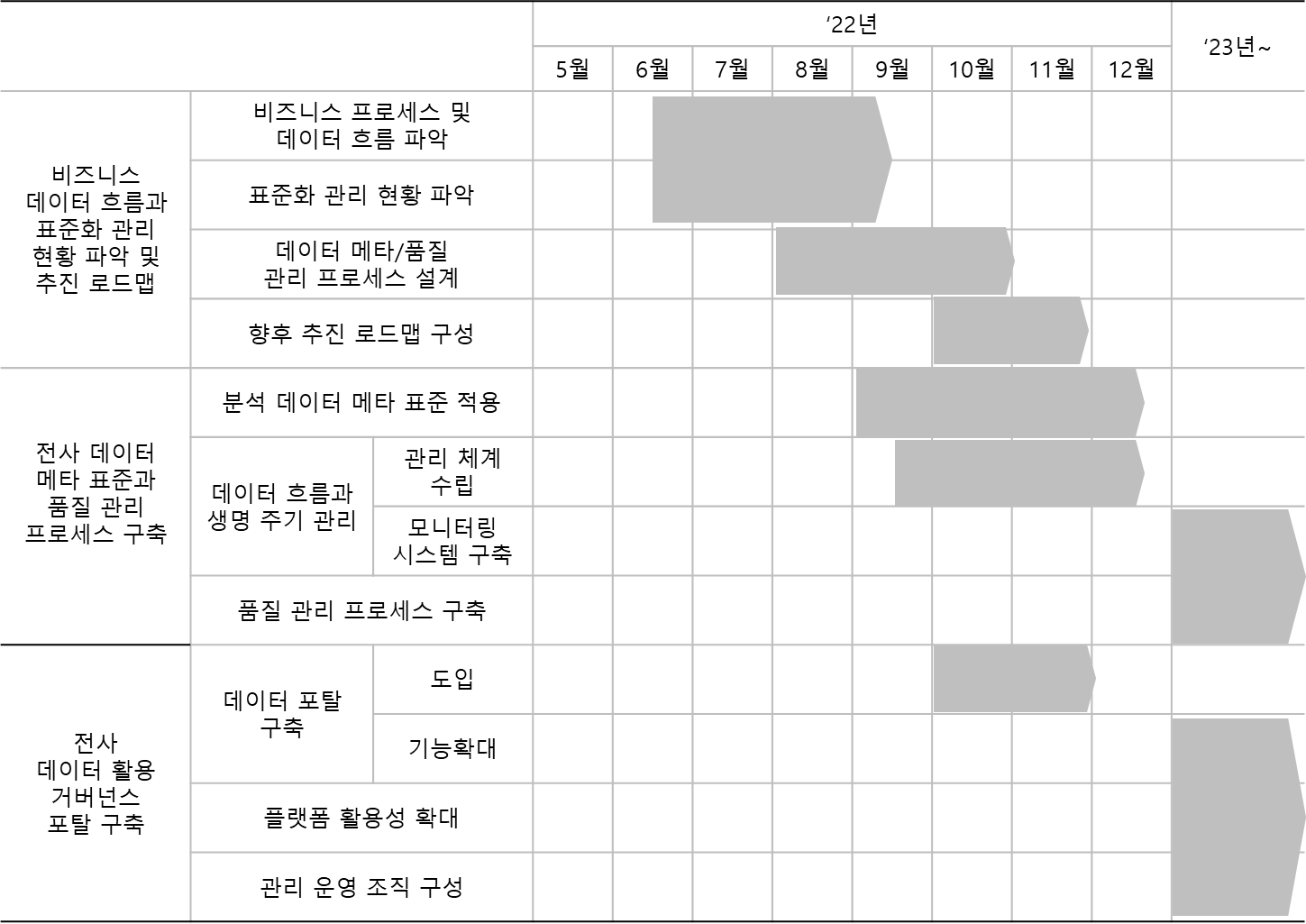


**[그림 1. 데이터 활용 거버넌스 구성]**

* **각 목표 업무에 대해서는 다음과 같이 추진 전략(방향성)을 설정하여 업무를 추진 하고자 함**
  + 1. 현행 “비즈니스 데이터 흐름과 표준화 관리 현황 파악 및 추진 로드맵” 구성
  + 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 확정 함
  + 더불어, 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 수립 함
  + 데이터 수집과 활용에 대한 보안 체계를 검토하고 (개인정보 보유 범위 등), 조직별 활용 데이터 접근 권한 체계 지침에 따라 구성 함  
    (데이터 활용 보안은 전사 보안 지침에 따름, 조직별 데이터 접근 권한은 데이터 거버넌스팀에서 정의)
  + 데이터 구성 변화 관리 프로세스를 수립하고 처리 방안에 대한 유관 부서와의 협의하여 정립 함
    1. 정보 분석 관점의 “데이터 메타 표준 구축과 품질 관리 프로세스 설계”
  + 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 구성
  + 각 사업부별로 관리되고 있는 데이터 속성 표준 코드를 통합하고 정제하여 정보 분석 활용에 적용할 수 있도록 코드 표준화와 원천(소스) 데이터 매핑 규칙을 시스템에 적용 함  
    ※ 상품/고객 데이터 표준화 포함
  + 데이터 품질 관리를 위하여, 수집하고 활용하는 데이터들의 메타 표준 준수 여부를 파악할 수 있도록 주기적인 품질 관리 자동화 수행과 모니터링 체계를 설계 함. 품질 진단 파일럿을 실행하여 진단 결과에 대한 개선 사항을 유관 부서에 공유하고 대응 조치 협업 프로세스를 구성함으로써 데이터의 품질을 지속적으로 확보할 체계를 마련 함
  + 원천(소스)부터 활용까지 데이터의 흐름을 정의하고, 이러한 흐름 정보를 시스템 내 구성하여 데이터 생명주기에 대한 모니터링 체계를 설계 함
    1. 데이터 표준 활용성 증대와 관리 운영 프로세스 위한 “전사 데이터 활용 거버넌스 포탈 구축”
  + 임직원들의 신속한 데이터 분석과 활용을 위하여 데이터 구성 요소 조회 및 검색 기능과, 데이터 흐름에 따른 데이터 속성의 매핑 규칙 조회 기능을 구성 함
  + 일반 사용자, 데이터 관리자(DA, DBA)의 권한 별 업무 처리 프로세스와 기능을 구성 하고 관리자가 데이터 구성 변화에 대응하기 위한 데이터 구성 변경 공유 기능을 적용 함
  + 데이터 품질 진단 대상을 확대하고, 품질 진단 결과 개선 적용 프로세스를 시스템에 적용 함
  + 데이터 활용 주체들의 분석 데이터 요청과 데이터 개선 사항 등의 요건 반영 프로세스를 시스템적으로 관리하여, 분석 업무 처리의 효율성을 확보 함
  + “전사 데이터 거버넌스 플랫폼 활용 가이드” 를 작성하고, 정보 분석과 활용이 필요한 유관 부서에 주기적으로 배포하고 교육을 진행 함  
    ※ 전사 데이터 거버넌스 플랫폼 활용 가이드 : 시스템의 활용 교육 뿐만 아니라 분석 데이터 변경 현황에 대한 공유도 포함

1. 구체 과제 및 해결 방안

* **‘II. 추진 방향” 에서 정의한 데이터 거버넌스 전략에 대하여, 아래의 [그림 1] 과 같이 상세 태스크를 도출하여 추진하고자 함**

****

**[그림 2. 데이터 거버넌스 과제 추진 일정 계획]**

|  |  |  |
| --- | --- | --- |
| 구 분 | 수행 범위 | 기대 효과 |
| 1단계  (22년 내) | 1. 비즈니스 데이터 흐름과 표준화 관리 현황 파악 및 추진 로드맵 1-1. 비즈니스 프로세스 및 데이터 흐름 파악 1-2. 표준화 관리 현황 파악 1-3. 향후 추진 로드맵 구성 2. 데이터 활용 거버넌스 메타 표준과 품질 관리 프로세스 구축 2-1. 활용 데이터 메타 표준 정의 2-2. 활용 데이터 품질 관리 프로세스 설계 2-3. 데이터 흐름과 생명 주기 분석  - 데이터 흐름과 생명 주기 관련 연계 아키텍처 구성 | 1) 분석 데이터 표준 수립으로 명확하고 일관성 있는 데이터 구성 환경 제공  2) 분석 속성(지표)의 명확한 의미 정의로 데이터 이해 증진  3) 분석 데이터 관리 프로세스 설계로 데이터 활용 효율성 향상 기반 마련 |
| 2단계  (23년 예정) | 1. 전사 데이터 메타 표준과 품질 관리 프로세스 구축 2-3. 데이터 흐름과 생명 주기 관리 구축  - 데이터 흐름과 생명 주기 관련 정보 수집  - 데이터 흐름과 생명 주기 모니터링 구축 2. 전사 데이터 활용 거버넌스 포탈 구축 3-1. 포탈 기능 구성 (사용자 관점) 3-2. 포탈 기능 구성 (관리자 관점) 3-3. 플랫폼 활용성 확대 3-4. 관리 운영 조직 구성 | 1) 주기적 품질 점검으로 분석 데이터 안정적 신뢰성 확보  2) 분석 데이터 획득과 활용 효율성 증대  3) 데이터 활용 처리 프로세스 적용  - 체계적인 데이터 관리/활용 체계 구성으로 업무 효율성 확보  - 분석 데이터 신속한 변화 대응 |

**[표 2. 단계별 추진 과제]**

1. 현행 “비즈니스 데이터 흐름과 표준화 관리 현황 파악 및 추진 로드맵” 구성

|  |  |
| --- | --- |
| **구체 과제** | **상세 태스크** |
| **비즈니스 프로세스 및 데이터 흐름 파악** | * **사업부별 비즈니스 데이터 이해**   + - * + **비즈니스 프로세스 파악 및 데이터 흐름 파악 (ER 분석 & 인터뷰)** * **분석 데이터 구성 파악**    + - * + **기존 정보 분석계(DW, 데이터레이크 등)에 구성한 데이터 구성 및 데이터 흐름 파악 (w/ 데이터서비스팀, 데이터레이크TF)** |
| **표준화 관리 현황 파악** | * **사업부별 데이터 표준 자료 수집**   + - * + **비즈니스 데이터 표준화 현황 관련 내부 자료 수집 및 분석, 데이터 구성 요소의 명확한 의미 분석** * **분석 데이터 표준 자료 수집**   + - * + **분석 활용 관점 데이터 표준화 현황 관련 내부 자료 수집 및  분석, 데이터 구성 요소의 명확한 의미 파악** |
| **향후 추진**  **로드맵 구성** | * **추진 로드맵과 상세 계획 수립**   + - * + **전사 데이터 거버넌스 활용 플랫폼 구축을 위한 추진 로드맵 수립과 이에 따른 상세 일정 계획 수립** * **데이터 거버넌스 보안 정책 수립**   + - * + **데이터 거버넌스 사용자별 조직별 데이터 접근 권한 정책 수립**         + **데이터 거버넌스 사용자별 조직별 업무 처리 권한 정책 수립** * **데이터 거버넌스 업무 프로세스 정책 수립**   + - * + **데이터 거버넌스 업무 정책 방향, 구성 조직, 업무 처리 프로세스 상세 정의(데이터 변경 관리, 데이터 수집/변경 요청)**         + **데이터 관리자(DA, DBA 등) 업무 처리 프로세스 상세 정의** |

1. 정보 분석 관점의 “전사 데이터 메타 표준과 품질 관리 프로세스 구축”

|  |  |
| --- | --- |
| **구체 과제** | **상세 태스크** |
| **분석 데이터**  **메타 표준 적용** | * **데이터 메타 표준 정의**   + - * + **정보 분석 데이터 거버넌스 메타 표준(단어/용어/도메인) 통합 및 정제**         + **표준 용어의 명확한 의미 정의** * **데이터 항목의 매핑 규칙 정의**   + - * + **원천(소스)에서 목표(타겟)으로의 데이터 추출/변환/적재(ETL) 과정에서 정의된 데이터의 매핑 규칙 정의**         + **원천(소스)에서 목표(타겟)으로의 데이터 처리 규칙 정의** |
| **데이터 품질 관리**  **프로세스 구축** | * **데이터 품질 진단 규칙(Rule) 정의**   + - * + **품질 진단 대상 데이터에 대한 품질 진단 규칙(도메인 규약, 필수값 여부, 표준코드 준수 여부, 비즈니스 규칙 등)** * **수집 데이터 품질 관리 시스템 구축**   + - * + **수집 데이터(예 : ODS)의 품질 진단 자동화 (정기/비정기) 및 진단 결과 모니터링 체계 구축 ※ ODS (Operational Data Store) : 운영 데이터 저장소**         + **수집 데이터 정합성 검증 및 진단 결과 개선 권고 프로세스 적용** * **분석 데이터 품질 진단 결과 공유**   + - * + **분석 데이터의 품질 진단 자동화 (정기/비정기) 및 진단 결과 모니터링 체계 구축**         + **분석 데이터 진단 결과 파악 및 개선 반영 요청 프로세스 적용** |
| **데이터 흐름과 생명 주기 관리** | * **데이터 흐름과 생명 주기 관련 정보 수집**   + - * + **정보 분석에 활용되는 데이터 흐름 (ETL, BI 등) 관련 솔루션의 데이터 매핑 정보와 사용 빈도 관련 정보 수집  ※ 정보계 솔루션(ETL, BI 등)과의 연동 개발 필요** * **데이터 흐름과 생명 주기 모니터링 구축**   + - * + **데이터의 수집부터 변환, 활용까지의 데이터 흐름을 모니터링하고 실제 활용 빈도를 주기적으로 파악하여 생명 주기 관리 함** |

1. 데이터 표준 활용성 증대와 관리 운영 프로세스 효율성을 위한 “전사 데이터 활용 거버넌스 포탈 구축”

|  |  |
| --- | --- |
| **구체 과제** | **상세 태스크** |
| **포탈 기능 구성**  **(사용자 관점)** | * **데이터 카탈로그 조회 구축**   + - * + **정보 분석에 필요한 데이터 메타 속성 조회(데이터 카탈로그)**         + **전사 활용 데이터 매핑 규칙 조회**         + **정보 분석을 위한 데이터 수집 및 변경 요청**         + **데이터 흐름 파악을 위한 화면 구성** * **데이터 요청 프로세스 구축**   + - * + **데이터 분석을 위한 데이터 수집과 구성 요건을 포탈을 통하여 요청하고 처리 과정을 확인하는 기능 구성** |
| **포탈 기능 구성**  **(관리자 관점)** | * **데이터 변화 관리 대응 프로세스 구축**   + - * + **수집(운영) 데이터 모델 구조 변경을 모니터링 하고 결과를 공유하여 즉각적인 대응이 가능한 프로세스 적용**         + **분석 데이터 모델 개체의 변경/추가/삭제에 따른 영향도를 파악하고 적용 절차에 따른 시스템 적용(DB 반영) 프로세스 구축 (메타 생성 및 데이터 스크립트 생성을 위한 데이터관리자(DA, DBA 등) 승인 절차 적용)** * **데이터 요청 대응 프로세스 구축**   + - * + **분석 데이터 수집과 추가 구성 요건에 대한 대응 부서 할당 및 처리 결과 반영** * **데이터 접근 권한 설정**   + - * + **조직별 사용자별 데이터 접근 권한 관리 화면 구성 (분석 데이터 물리적 모델 구성 시 참고하여 접근 권한 부여)** |
| **플랫폼**  **활용성 확대** | * **플랫폼 활용성 전략 수립**   + - * + **플랫폼 활용 가이드 작성과 전사 주기적 배포**         + **시스템 사용법 정기적인 교육 계획 수립 후 추진** |
| **관리 운영**  **조직 구성** | * **전사 데이터 거버넌스 운영 조직 구성**   + - * + **표준 메타, 표준 지표 정의, 모델 구성 모니터링과 물리 생성, 활용 가이드 작성 및 교육에 필요한 데이터 거버넌스 운영 조직 구성**         + **데이터 변경 관리 적용을 위한 유관 부서 협업 프로세스 정립** |

* + 1) 2) 과제 수행을 위한 1단계 예상 기간 및 투입 인원 계획은 다음과 같음
* **기간 : 2022년 07월 ~ 2022년 12월 (1단계, 약 6개월)**
* **비용 : 총합 약 7.5억 (VAT 별도, 내부 인프라 활용)  
  (세부 항목 – 외부 용역(메타표준 적용 및 관리 정책수립) : 약 4.5억, 솔루션 도입 : 약 3.0억)**
* **인원 : (내부 2명) 20 M/M (외부 5명) 25 M/M**
* **범위 : 현행 비즈니스 기반 데이터 표준 파악, 현행 분석 데이터 표준 파악, 분석 데이터 메타 표준 정의 및 적용, 분석 데이터 관리/보안/활용 정책 정의, 데이터 품질/흐름 관리 프로세스 설계, 데이터 포탈 상세 구축 범위 도출 및 추진 로드맵 수립**

1. 예상 결과물

* **본 과제 수행을 통한 기대 효과는 다음과 같음**
* **본 전략 과제 수행을 통하여 데이터 활용 관점에서 분석 지표 의미의 명확성을 제공함과 동시에 검증되고 일관된 데이터 품질을 확보하여 조직 구성원들의 데이터에 대한 이해도와 활용 능력을 증대할 수 있음. 더불어 데이터 활용 시 신속한 업무 처리 프로세스를 적용하여, 구성원들의 전략적 사업 기획과 추진을 위한 소요시간(Lead Time)을 단축시키고, DT, DX에 기반한 미래 사업 발굴과 수행 성공율을 높여 당사의 시장 경쟁력과 인지도를 강화할 수 있는 토대를 마련 할 수 있음**
* **최종 결과물로는 다음과 같은 산출물을 작성하고자 함. (22년 12월 기준)**
  + - 사업부 비즈니스별 현황 분석 결과서
      * 사업부별 데이터 흐름도 및 현황 분석서
      * 표준화 관리 현황 분석서
    - 정보계 대상 “데이터 활용 거버넌스 추진 로드맵”
      * 추진 단계별 상세 일정 계획(WBS)
    - 정보 분석 영역에 대한 “데이터 메타 표준 및 품질 관리”
      * 전사 데이터 메타 표준 정의서
      * 데이터 품질 관리 프로세스 흐름도
      * 전사 통합된 데이터 관리 정책
    - 정보 분석 영역에 대한 “전사 통합 데이터 포탈”
      * 요구 사항 정의서
      * 데이터 거버넌스 포탈 업무 프로세스 설계서
      * 데이터 거버넌스 포탈 화면 설계서
      * 관리 시스템
        + 데이터 품질 관리 시스템
        + 데이터 메타 관리 시스템

- 끝 –