

ESG공시 자동화 시스템 기획안

팀 ConanAI

팀원 천준영, 김하늘

0. 제도적 배경 및 개발 방향

국내 ESG 공시제도는 금융위원회가 정책 수립과 규제 도입을 주도하고 있으며, 공시의 실제 집행 및 기술 시스템 운영은 금융감독원이 담당하고 있다.

현재 ESG 공시 의무화 전환과 관련하여, 금융위원회가 공식적으로 제시한 방향은 다음과 같다.

첫째, ERSR 및 ISSB 기준을 기반으로 한 KSSB 공시기준 가이드라인을 제정하고 있으며,
둘째, 기후 관련 공시를 우선 의무화하고, 향후 기타 ESG 항목을 단계적으로 의무화 대상에 포함할 계획이다.
셋째, 보고서의 제출 형식(XBRL 또는 DSD 등)에 대해서는 아직 구체적인 기술 형식이 제도화되지 않은 상태이다.

한편, EU(CSRD), 일본(ISSB) 등 주요국이 ESG 공시를 법정 공시 체계로 전환하고 있는 상황을 고려할 때, 한국 또한 자본시장법 개정을 통한 법제화 가능성이 높다. 만약 ESG 공시가 자본시장법상 공시로 편입된다면, 공시는 기존의 전자공시시스템(DART)을 통해 이루어지게 되며, 현행 전자공시 관행에 비추어 볼 때 XBRL 및 DSD 형식의 구조화된 제출 방식이 채택될 가능성이 높다.

이에 따라 본 프로젝트는 향후 제도적 변화에 선제적으로 대응할 수 있도록, ESG 보고서를 XBRL 및 DSD 형식으로 자동 생성하는 시스템 구조로 설계되었다. 특히, 금융위원회가 1차 공시 대상으로 명시한 '기후 관련 공시'를 중심으로 개발을 시작하며, 보고서 구조는 ERSR E1, ISSB S2, EU Taxonomy 기준에 기반하여 구성된다.

정량 정보의 경우, XBRL 형식으로 자동 생성할 예정이며, 현재 금융감독원이 ESG 전용 XBRL taxonomy를 제공하고 있지 않은 상황을 고려해, EU Taxonomy 기반의 XBRL 구조를 내부 참조 스키마로 활용할 계획이다. 단, 이는 비공식 구조이므로, 실제 DART 제출에는 제약이 따를 수 있음을 사전에 인지하고 있다.

정성 정보는 DSD 형식의 보고서 내 서술형 항목으로 구성되며, .dsd 파일 자체는 금감원 제공 태그 및 해시 검증이 요구되기 때문에, 파일 생성은 보류하고 우선적으로 DSD 서술부 콘텐츠만 자동 생성하는 방식으로 접근한다. 이는 향후 DSD 기반 ESG 공시가 정식 도입될 경우를 대비한 구조적 대응 전략이다.

최종적으로 본 시스템은 ESG 공시 항목을 다음과 같이 구분하여 자동화한다.

- 정량지표는 XBRL 태깅을 통해 구조화된 XML 파일로 생성
- 정성지표는 DSD 보고서 서술부에 포함될 내용을 자동 생성

이때 적용되는 공시 프레임워크는 다음 두 가지 기준에 기반한다:

- ERSR/ISSB의 '지배구조-전략-위험관리-지표 및 목표'
 - EU Taxonomy의 Turnover, CapEx, OpEx 세 가지 정량 지표 구조
-

1. 프로젝트명

ESG 전자공시 자동화 시스템: XBRL/DSD 기반 보고서 생성 플랫폼

2. 프로젝트 개요

본 프로젝트는 금융위원회 주도 하에 예정된 ESG 의무공시 제도 도입에 대응하여, 기업이 ESG 데이터를 입력하면 이를 XBRL 및 DSD 형식의 전자보고서를 자동 생성하는 기능을 지원하며, 사용자가 최종 검토 후 확정할 수 있는 보조 시스템을 포함한다.

프로토타입은 기후 관련 공시 정보(ESRS E1 / ISSB S2 기준)를 중심으로 설계되며, 향후 타 ESG 주제로 확장 가능하도록 구조화된다.

3. 개발 목적

- 제도 대응: ESG 의무공시가 자본시장법 하의 전자공시 체계(DART)로 편입될 가능성에 대비
 - 형식 표준화: XBRL(정량) 및 DSD(정성) 형식을 구조화하여 제출 요건 충족 준비
 - 자동화 효율화: 비정형적 ESG 데이터의 보고서화 과정을 자동화하여, 기업의 공시 비용 및 오류율 감소
 - 선도사례 확보: ESRS XBRL taxonomy 구조를 참조하여 국내 최초 자동화 사례 구현
-

4. 전체 시스템 구조 요약

1) DSD 서술문장 생성

① 기후관련 데이터 입력

↓

② 적합성 계산

ESRS E1, ISSB S2 -> 요구사항을 체크리스트화 하여 각각에 대응되는 정보가 공시되었는지 확인

EU taxonomy -> 기술심사기준+DNSH를 기반으로 EU Taxonomy 적합 여부 판단.

↓

③ DSD보고서 작성

AI언어모델을 통해 ESRS/ISSB의 '지배구조-전략-위험관리-지표 및 목표(G/S/R/M) 및 EU

Taxonomy의 Turnover, CapEx, OpEx 3요소를 기반으로 템플릿 기반의 서술형 문장 작성

2) XBRL 보고서 생성

① 기후관련 정량 데이터 입력

↓

② 적합성 계산

ESRS E1, ISSB S2 -> 위와 동일

EU taxonomy -> 위와 동일



③ 정량 데이터 항목을 ESRS XBRL taxonomy 태그와 매핑



④ XML구조에 맞춰서 XBRL파일 생성

5. 기능별 자동화 구성요소

① 기후 관련 데이터 입력 (입력 구조 정형화)

- 기능 설명

ESG 공시에 필요한 입력 정보를 사용자가 직접 입력하도록 하되,
국제 기준(ESRS E1, ISSB S2, EU Taxonomy)에서 요구하는 사항을 기반으로
체크리스트 및 선택형/서술형 입력 필드로 구성
- 기술 구현 방안
 - 입력 템플릿: HTML/React 기반 입력 폼
 - 보조 기능:
 - 항목별 설명 톨팁
 - 예시 답안 보기
 - 중요도 필터링(필수/선택)
 - 구조 분류: '지배구조', '전략', '위험관리', '지표 및 목표'
 - 편의 기능: 저장 후 나중에 입력 가능
- 자동화 범위
 - 자동완성
 - 예시 문장 팝업 제공
 - PDF 파싱 모듈(보조기능, OCR 미포함)

② 적합성 평가 및 사전 진단 (Validation Engine)

(1) ESRS/ISSB 평가

- 구현 로직
 - 사전 정의된 요구사항과 입력값 매핑 (Rule-based)
 - 예: "전략 항목에 '시나리오 분석' 언급 있음?" → Boolean 체크
 - 문장 유사도 평가
 - 기존 TCFD/ESG 보고서 문장을 SBERT 또는 MiniLM으로 벡터화
 - 사용자 입력 문장과 유사 문장 비교 → '충분한 내용' 추정
 - 유사도 기반 평가 결과는 근거 문장을 시각화로 제공

- Human-in-the-loop 보완
 - 사용자 수정 허용
 - 관리자 검토 기능 추가 가능

(2) EU Taxonomy 평가

- 구현 로직
 - 활동명 → NACE 코드 후보 매핑
 - keyword-based search + 후보 제안 → 사용자 최종 선택
 - DNSH 및 기술심사 기준 적용
 - Annex I/II 체크리스트화
 - 입력 조건 기반 '충족/불충족/미확인' 상태 표시
 - 정량 계산
 - taxonomy-eligible 매출/CapEx/OpEx 입력값 정렬
 - 총 매출 대비 비중 자동 계산, 표로 정리

③ 보고서 문장 생성 및 DSD 템플릿 작성

(1) ESRs/ISSB 서술 파트

- 작성 방식
 - G/S/R/M 구조별 템플릿 기반
 - 대표 문장 프레임 고정
 - 예: “당사는 [위험명]에 대해 [전략]을 수립하였으며, [지배구조 기구]에서 관리 중입니다.”
 - 사용자 입력값 → 변수 삽입
- LLM 활용
 - GPT 또는 KoGPT 활용 (prompt-based)
 - few-shot 예시로 문체 통일
 - fine-tuning은 예산 제약으로 미 실시

(2) EU Taxonomy 서술 파트

- 작성 방식
 - 수치 기반 고정 템플릿
 - 예: “당사의 CapEx 중 [XX]%는 EU Taxonomy 기준에 부합하는 활동에 사용되었습니다.”
- 출력물
 - HTML 또는 Markdown 형식의 DSD 보고서
 - <section>, <table> 구조로 서술부/수치부 분리 출력

④ XBRL 태깅 및 구조화

- 기능 개요
공시 적합성 평가를 통과한 정량 정보를 기반으로, **ESRS XBRL taxonomy**의 태그와 자동 매핑하고 **XML** 구조로 변환
- 자동화 범위
 - 정량 항목별 태그 후보 자동 매핑
 - 단위(unit), 문맥(context), 기준일 고려
 - context, unit, schemaRef 포함 XML 요소 자동 생성
 - 조건부 필터링 기반 태그 확정
 - 반복 항목 구조 자동 출력
 - XBRL 기술 사양 준수

⑤ 유효성 검사 및 결과 출력

- 기능 개요
생성된 **XBRL** 파일은 공식 **XSD** 스키마에 따라 유효성 검사를 거치고, 오류 발생 시 상세 리포트를 통해 피드백
- 자동화 범위
 - lxml 또는 arelle 기반 유효성 검사 자동 수행
 - 오류 리포트 생성
 - 태그 누락
 - 단위 오류
 - context 불일치
 - 경미 오류에 대해 자동 수정 제안
 - 출력물 패키징
 - XBRL XML 파일
 - CSV/JSON 요약본
 - 미리보기 UI 연동

6. 기술스택

분류	주요 기술 및 라이브러리
백엔드 API	Python, FastAPI, Uvicorn, Pydantic
프론트엔드	Next.js 15, React 19, Zustand, Tailwind CSS

데이터 입력 및 분석	Pandas, NumPy, openpyxl
보고서 문장 생성 (DSD)	OpenAI GPT, KoGPT, Prompt Engineering
텍스트 유사도 평가	SentenceTransformer (SBERT, MiniLM)
문서 구조 파싱	PyMuPDF, PDFMiner, Markdown Renderer
XBRL 생성 및 처리	lxml, xmlschema, custom ESRS taxonomy (XSD), arelle
유효성 검사 및 오류 리포트	lxml.XMLSchema, arelleCmdLine
파일 출력 및 미리보기	Markdown → HTML 변환, FastAPI Static File Response
배포 및 환경관리	Docker, Docker Compose, python-dotenv