

별첨1-5 [제출물 양식] 활용데이터 목록 (Track1)

연 번	데이터 개요	데이터 구분 (공공/개인/ AI 합성)	데이터 수집경로 (AI 합성 방법)
1	<ul style="list-style-type: none"> NCBI GEO (GSE134900) - 셀리악병(Celiac Disease) 환자 및 정상 대조군의 십이지장 생검 조직 RNA-seq 전사체 데이터 (\$n=96\$) 	공공	<ul style="list-style-type: none"> URL: https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE134900 (Perplexity를 통해 원본 출처 확보 및 무결성 검증 완료)
2	<ul style="list-style-type: none"> Synthetic Celiac Transcriptomics - GSE134900의 잠재 분포를 학습하여 생성한 고품질 가상 환자 전사체 데이터 (n=1,000) 	AI 합성	<ul style="list-style-type: none"> 생성 모델: VAE (Variational Autoencoder) with Sampling Temperature Scaling (T=2.0) 구현 환경: PyTorch / Cursor (Claude 4.5 Opus High) (상세 알고리즘은 연구보고서 3. Methodology 참조)
3
4
5