

AI를 이용한 통계자료 기반 보도자료 자동작성 시스템 개발

1. 문제 정의

- (현황) 통계청은 66종의 승인통계를 작성하고 있고, 매월 월간통계를 포함하여 20건 내외의 보도자료 형태로 공표 중
 - 수집된 자료를 DB, 엑셀 등에 정리한 후, 그 집계 결과를 활용하여 표, 그래프, 설명 문구 등을 정형화된 보도자료로 작성
- (문제점) 동일한 패턴의 통계표, 그래프, 설명 문구 작성 반복 작업으로 비효율 발생

2. 제안 프로젝트

- (내용) 엑셀 및 KOSIS API 기반 자료 연계를 통한 지역경제동향^{분기} 보도자료 자동 작성 시스템 개발
 - 데이터 전처리: 엑셀 자료 로딩, KOSIS API 연계 자료 수집 체계 설계 및 각 수치에 붙은 메타 정보(단위, 지역 등) 확인
 - 표, 그래프 자동 생성: 수치 자료 연계를 통해 통계표 생성, 파이썬 등 활용하여 시계열 및 지역별 그래프 작성
 - 설명 문구 작성: 보도자료 내 “▲▲지역 생산은 전년동분기대비 ○% 증가” 등 설명 문구 자연어 생성 규칙과 LLM을 활용하여 자동 생성
 - 보도자료 템플릿 구성 및 모듈화
 - 정확성 검증 시스템 마련 및 매뉴얼 작성
- (제공 데이터) 지역경제동향 보도자료 및 엑셀 기초자료 등
- (목표) 수작업 및 작업시간을 최소화한, 정확성 있는 보도자료 작성
 - 지역경제동향 보도자료 외 타 통계에도 적용 가능한 AI 활용 보도자료 생성 모델 개발이 최종 목표

3. 요구 결과물

- 보도자료 자동 생성 결과물 샘플(hwp, pdf 등)
- 자동화 코드 패키지, 엑셀 및 KOSIS API 연동 모듈 등
- 최종결과 보고서 및 매뉴얼

4. 프로젝트 관련 필요 역량

- (통계지식) 지역경제 지표 해석 및 통계청 보도자료 작성법 이해
- (프로그래밍) Python(pandas, matplotlib, openpyxl), R 활용 가능
- (데이터 처리) API연동, 엑셀 자동화 등
- (자연어 생성) 기본적인 텍스트 자동 생성 원리 이해

5. 기타

- (근무공간) 나라샘도서관^{서울 강남}에서 근무
 - 통계청^{대전} 담당자와 비대면 및 대면^{주1회} 이상 도메인 지식 전수 예정
- (간담회 운영) 통계청 공공데이터 활용 및 AI 업무 관련 교육 및 담당자 간담회 진행