

# 한국어 문서 추출요약 AI 경진대회

## 주제

다양한 주제의 한국어 원문으로부터 추출요약문을 도출해낼 수 있도록 인공지능을 개발

## 배경

다양한 분야에서 뉴스기사 요약, 법률분석 요약, 사업보고서 요약 등 핵심내용을 신속하고 정확하게 파악할 수 있는 AI 요약기술을 필요로 하고 있음

## 과제

원문 텍스트에서 3개의 중요한 문장을 선택하는 추출요약 AI 개발

## 데이터

1. train.jsonl - 학습에 사용 할 데이터셋
  - media : 기사 미디어
  - id : 각 데이터 고유 번호
  - article\_original : 전체 기사 내용, 문장별로 split되어 있음
  - abstractive : 사람이 생성한 요약문
  - extractive : 사람이 추출한 요약문 3개의 index
2. extractive\_test.jsonl - 추론할 데이터셋
  - media : 기사 미디어
  - id : 각 데이터 고유 번호
  - article\_original : 전체 기사 내용, 문장별로 split되어 있음

## 일정

데이터공개		대회기간		코드제출		평가기간		결과 발표
11.04	➡	11.11 ~ 12.09	➡	12.13	➡	12.14 ~ 12.20	➡	12.21