

# DATA·AI 분석 경진대회 추진계획(안)

한국과학기술정보연구원

- (추진배경) 인공지능 데이터 및 기술을 기반으로 사회·과학 문제를 해결하고 국내 인공지능 생태계 저변을 확대



## □ (행사명) DATA·AI 분석 경진대회

- (주최/주관) 국가과학기술연구회, 국회도서관, 대전광역시, 해양경찰청, 대전중구청\* / KISTI
  - \* 문제제출 기관 요청시 주최기관 추가 예정
- (후원) 한국콘텐츠학회
- (부분) 공모를 통해 문제발굴 후 문제해결 순차 진행
  - (문제발굴 부문) 자유주제 발굴과 지정주제 발굴로 구성
    - AI 기술 및 데이터 분석을 통해 해결이 필요한 다양한 사회·과학 문제 및 데이터 발굴
    - 수요기관이 지정한 주제 내에서 문제(아이디어)를 내거나, 자유 주제로 문제 제시 가능
    - 평가를 통해 선정된 문제는 문제발굴상을 수여하며, 문제해결 단계로 진행
  - (문제해결 부문) 공모발굴 문제와 정책지정 문제로 구성
    - 문제발굴 단계에서 선정된 공모 발굴문제 및 수요기관이 제시한 정책지정 문제를 해결하는 높은 성능의 인공지능 모델 개발
    - 문제별 제출 결과 중 우수한 해결책들은 우수 모델 후보로 선정
    - 발표평가를 통해 문제별 최우수 해결책은 수요기관이 시상하는 우수문제해결상을 수여하며, 전체 우수 모델후보들을 대상으로 우수 모델들은 특별상을 추가 수여
- (참가자격)
  - (문제발굴 부문) 사회·과학 당면 문제 해결이 필요한 기관이나 개인
  - (문제해결 부문) AI 데이터 분석·활용에 관심이 있는 개인이나 팀

□ (행사일정) ‘24. 7. 10. ~ 11. 20. (변경 가능)

진행 내용	일정	비고
문제발굴	7. 10. ~ 8. 16. (경진대회 기간 이후 상시 진행 예정)	경진대회 플랫폼
채택 문제 공개	8. 23.	경진대회 플랫폼
문제해결	8. 23. ~ 10. 25.	경진대회 플랫폼
(예선) 문제별 후보 선정*	11. 7 ~ 11. 8.	온라인 평가
(결선) 우수 모델 선정	11. 13. (추후 확정)	발표 평가
시상식	11. 20. (추후 확정)	미래연구정보포럼

\* 제출 결과 중 수요기관에서 해결방안으로 활용 가능한 후보 선정 (문제별 2개 이내)

□ (시상내역) 총상금 2,400만원 / 15개 팀 선발\*

부문	상장	수상자(팀)	시상 내역	상금
문제발굴	우수문제 발굴상	4인	• KISTI 원장상	각 50만원
문제해결	우수문제 해결상	6팀	• 해양경찰청장상 • 국회도서관장상 • 대전중구청장상 • 수요기관 추가예정	각 200만원
특별상**	대상	1팀	• 국회의장상	400만원
	최우수상	2팀	• 연구회이사장상 • 대전광역시장상	각 200만원
	우수상	2팀	• KISTI 원장상	각 100만원

\* 발굴되는 문제 수 및 해결 방법의 우수성에 따라 수상팀의 수는 변동될 수 있으며, 중복 수상이 가능함

\*\* 문제해결 부문 우수문제모델 후보를 대상으로 결선 발표평가를 통해 선정

**aida** 인공지능 데이터 서비스  
 aida.kisti.re.kr  
 https://aida.kisti.re.kr/competition



# AI DATA

## 2024 DATA·AI 분석 경진대회

인공지능 데이터 및 기술을 기반으로  
 사회·과학 문제를 해결하여  
 국내 인공지능 생태계 저변을 확대



기간 **07. 10(수) ~ 11. 20(수)**

● 문제발굴 부문 ●

다양한 사회·과학 문제를 발굴하고 관련 데이터를 공개하여  
 다양한 연구자의 해결 참여 유도

**참가자격** 사회·과학 당면 문제 해결이 필요한 기관이나 개인

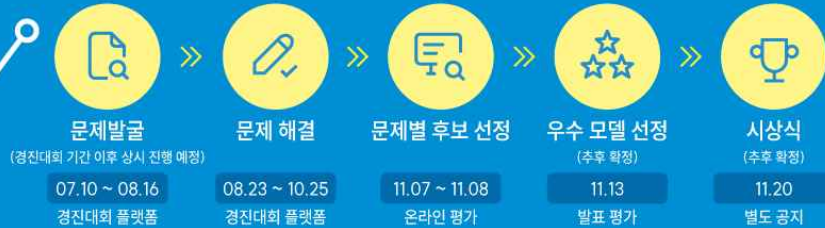
● 문제해결 부문 ●

제시된 사회·과학 문제 해결을 위한  
 높은 성능의 인공지능 모델 개발

**참가자격** AI 데이터 분석·활용에 관심이 있는 개인이나 팀



### 대회일정



### 시상 내역

총상금 **2,400** 만원

부문	구분	수상자(팀)	시상 내역	상금
문제발굴	우수문제 발굴상	4인	· KISTI 원장상	각 50만원
문제해결	우수문제 해결상	6팀	· 해양경찰청장상	각 200만원
			· 국회도서관장상	
특별상	대상	1팀	· 국회의장상	400만원
	최우수상	2팀	· 연구회이사장상	각 200만원
	우수상	2팀	· 대전광역시장상	
			· KISTI 원장상	각 100만원

· 발굴되는 문제 수 및 해결 방법의 우수성에 따라 수상팀의 수는 변동될 수 있음

주최/주관 **nst** 국가과학기술연구회 **국회도서관** **대전광역시** **해양경찰청** **대전중구청** / **KISTI** 한국과학기술정보연구원  
 후원 **KOON** (사)한국콘텐츠학회

**문의처** 경진대회 운영사무국  
 aidatacon@gmail.com

연구계, 기업계, 지역사회 등 다양한 영역에서 인공지능 기술 기반으로 해결하고자 하는 문제를 데이터와 함께 제시

**1. 문제 개요**

- 문제 제목
- 배경 및 목적
- 문제 설명

**2. 목표**

- 문제 해결 목표 성능 또는 기대 결과물에 대한 설명
- 평가 방법 및 설명

**3. 데이터 설명**

- 데이터명
- 유형, 포맷, 용량/건수 등
- 문제해결을 위해 필요한 설명 (데이터 사전)

\* 데이터 사전 예시 (붉은색은 필수)

속성	설명	데이터 타입	값 범위 혹은 범주	참고
파일명	사진 파일명	String		
국적	어선 소속 국적	String	중국, 한국	
선종	어선의 종류	String	타망, 범장망,...	
조업여부	조업중인지 이동중인지 여부	String	조업, 이동...	사진 내 파도와 그물 형태로 식별 가능
좌표	선박의 위치			사진 좌측 하단에 경도, 위도 순 표시

\* ... 표시는 분류 대상이 되는 모든 레이블 표기

○ 기타 참고 데이터에 대한 설명

예시) 공공데이터 포털의 API, 한국의 조업 영역 등

#### 4. 데이터 샘플

\* 문제해결의 아이디어를 도출하는데 필요한 정도의 샘플 데이터를 우선 제출한 후, 문제가 채택된 이후 문제 해결을 위한 충분한 데이터를 공개

○ 데이터 사전의 형식으로 데이터셋이 구성되어야 함

#### 5. 제출 형식

○ 참가자로부터 받을 결과 파일의 형식

\* 결과 파일 예시

파일명	국적	선종	불법 조업 여부
take1.jpg	중국	범장망	불법
take2.jpg	한국	안강망	합법

지역 경제 활성화 방안 모색을 위한  
성심당 방문객의 이동경로 및 소비패턴 분석

수요기관 : 대전광역시 중구청

**문제설명**

**문제 목적**

성심당 방문객의 이동경로 및 소비패턴 분석

**문제 배경**

성심당의 인기로 타 지역 방문객들이 증가함에 따라 이와 연계한 지역 경제 및 관광 활성화 정책 마련 요구

kbs 뉴스, 2024.5.13. 「성심당 매년 천만 명 찾지만... “빵만 사고 떠난다”」

- “대전에 대해 잘 모른다”, “지역 관광 홍보 필요”

kbs 뉴스, 2024.5.14. 「‘연간 천만 명’ 성심당 손님 잡아라... ‘빵빵한’ 효과 기대」

- [박삼화/성심당 상무이사 : "90% 가까이가 타 지역 분들입니다. 타 지역 분들의 분포를 한번 보면 수도권 분들이 일단은 가장 많고요.")]
- 인근 상인들의 적극 유치(영수증 가져오면 물건 값을 할인해 주고, 빵을 가져오면 식기를 내주는 카페 등) // 상생협력 「으능이랑 성심이랑」

**문제 해결 일정**

2024-07-01 ~ 2024-09-30

**결과 평가 방법**

정성평가 (창의성, 우수성, 효과성)

**데이터 제공**

- 샘플 데이터만 공개되며, 참가신청서 및 해결제안서를 제출한 후, 심사 및 온라인 설명회 이후 승인된 개인(팀)에게 전체 데이터 제공
- 샘플 데이터에 포함된 유동인구 및 카드매출 데이터는 난수로 합성된 자료이며, 참가 승인팀에게 원자료 제공 예정
- 주어진 데이터 이외에 다른 공개 데이터를 활용하여 목적을 달성할 수 있는 방법 제시 가능
- 제공되는 전체 데이터는 이번 경진대회 이외의 목적으로 사용, 유통, 판매 등을 절대 할 수 없음

## 결과 제출

다음 파일들을 zip으로 압축하여 제출

- . 자유양식으로 hwp 혹은 doc 으로 문제해결방법에 대한 설명 파일
- . 성심당 방문객의 이동경로 및 소비패턴 분석 결과는 연령대, 성별, 방문일자, 방문지역, 업종, 매출액 등을 포함하며 csv 등 재가공 가능한 파일 형태로 할 것
- . 예측 혹은 분석 코드 및 예측 모델

## 데이터설명

### 유동인구

- 설명: 관내 일별 유동인구 데이터
- 생산/수집: 통신사 제공
- 건수: 약 5,000,000건
- 형태: csv
- 데이터 정의: 첨부파일 참고 (유동인구정의서)

### 카드매출 (Cell Id별/업종별 가맹점 실적정보)

- 설명: Cell Id별 업종별 가맹점 실적정보
- 생산/수집: 카드사 제공
- 건수: 약 50,000건
- 형태: csv
- 데이터 정의: 첨부파일 참고 (카드매출정의서, 국토지리정보원\_국토조사(격자)500M)

### 카드매출 (지역별 업종 유입고객 및 상권특성 정보)

- 설명: 지역별 업종 유입고객 및 상권특성 정보
- 생산/수집: 카드사 제공
- 건수: 약 400,000건
- 형태: csv
- 데이터 정의: 첨부파일 참고 (카드매출정의서, 국토지리정보원\_국토조사(격자)500M)

### 상가(상권)정보

- 설명: 공공데이터 포털  
소상공인시장진흥공단\_상가(상권)정보(<https://www.data.go.kr/data/15083033/fileData.do>)
- 생산/수집: 소상공인시장진흥공단 수집(공공데이터 포털 게시)
- 건수: 77,393건
- 형태: csv
- 데이터 정의: 첨부파일 참고 (상가(상권)정보\_데이터코드, 상가(상권)정보\_업종분류)
- 데이터 다운로드 : 공공데이터포털