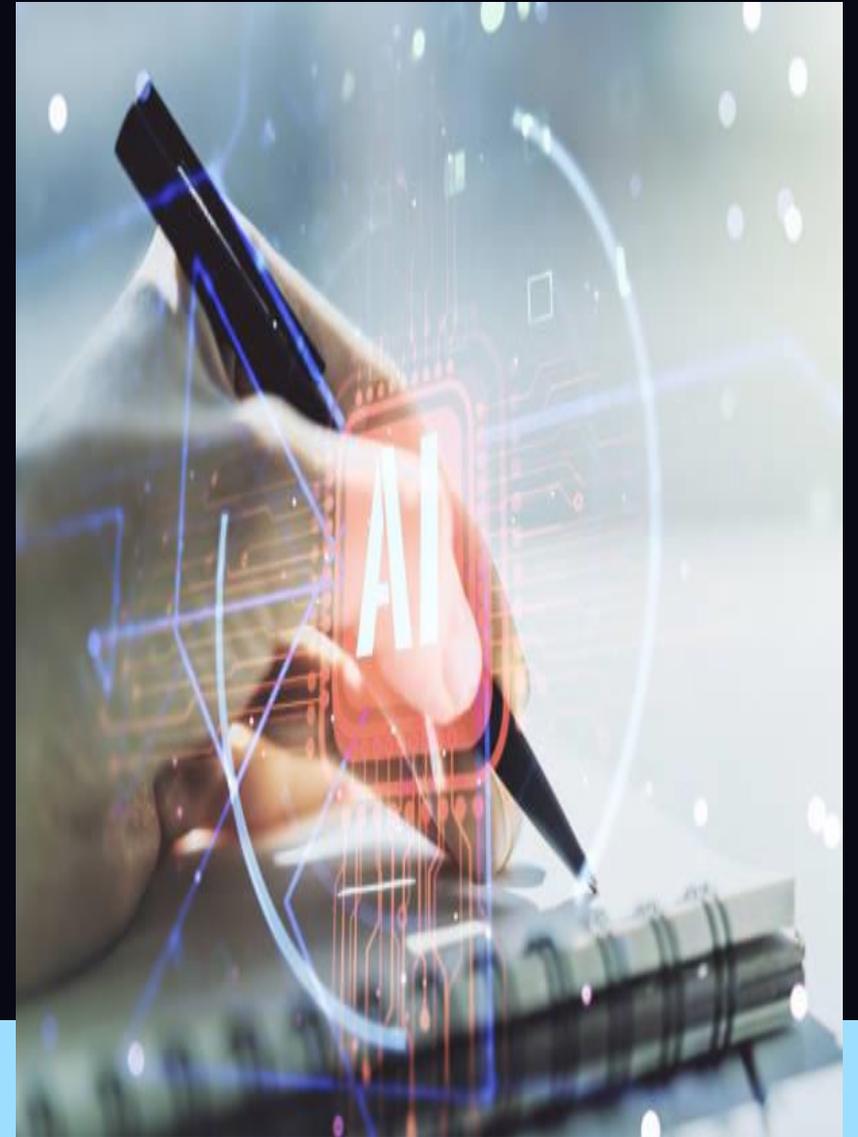


<‘인공지능 활용 국어능력 진단체계 개발 협력’ 협약 대학 연수회> 참가 보고

# 인공지능 활용 국어능력 진단체계 개발 사업 소개 및 글쓰기 자료 수집 방안 논의



2022.12.17.

박보연 (인천대 기초교육원)

# 목차

## 1. '인공지능 활용 국어능력 진단체계 개발' 사업 소개

- 가. 사업 개요 및 진행 상황
- 나. 국립국어원과 각 대학의 역할

## 2. 국어능력 진단체계 개발 기초 연구 결과

- 가. 연구 개요 및 국내외 진단체계 운영 실태
- 나. 문항 개발 및 평가 실행

## 3. '글쓰기 자료' 수집 방안 논의

- 가. '글쓰기 자료' 성격
- 나. '글쓰기 자료' 수집 분량 및 기한
- 다. '글쓰기 자료' 수집 가안
- 라. '글쓰기 자료' 수집 방안 논의





# 1. '인공지능 활용 국어능력 진단체계 개발' 사업 소개

## 가. 사업 개요 및 진행 상황

### ■ 사업 개요

#### · 사업 방향

공신력 있는 진단지표를 기반으로 대규모 글쓰기 자료 구축 및 인공지능 활용 국어능력 진단체계\* 운영으로 국민의 국어능력 향상

✓ 인공지능 활용 국어능력 진단체계:  
글의 내용, 조직, 표현을 분석하여 평가 전문가와 80% 이상의 일치도를 보이는 인공지능 글쓰기 채점 기술 개발을 목표로 함.



# 1. '인공지능 활용 국어능력 진단체계 개발' 사업 소개

## 가. 사업 개요 및 진행 상황

### ■ 사업 개요

#### · 기술적 차원의 목표

- 자동 채점을 위한 인공지능 모델 개발 목표
- 목표 모델 개발을 위한 기술적 과제
  - : 채점 지표별(내용, 조직, 표현) 자동평가 인공지능 모델 개발 목표
  - : 다양한 평가 문항 대상 자동평가 일반화 기술 개발



# 1. '인공지능 활용 국어능력 진단체계 개발' 사업 소개

## 가. 사업 개요 및 진행 상황

### ■ 사업 개요

#### · 활용적 차원의 목표

1단계 목표: 대규모 국어 쓰기 능력 평가에서 자동 채점의 활용

2단계 목표: 쓰기 능력 공인 인증시험 및 모의테스트 채점 서비스

3단계 목표: 쓰기 학습을 위한 쓰기 결과물 점검 및 피드백 서비스



# 1. '인공지능 활용 국어능력 진단체계 개발' 사업 소개

## 가. 사업 개요 및 진행 상황

### ■ 사업 개요

- 2022년 주요 사업

인공지능 활용 국어능력 진단체계 개발 기초 연구  
(서울대 국어교육연구소)

→ 인공지능 학습용 쓰기 자료 확보를 위한 원천 자료의 시범적 구축



# 1. '인공지능 활용 국어능력 진단체계 개발' 사업 소개

## 가. 사업 개요 및 진행 상황

### ■ 사업 개요

#### • 2023년 주요 사업

- ✓ 국어능력 진단체계 개발 협의체 구성(연 2회 진행 예정)
- ✓ 평가 지표의 고도화
- ✓ 글쓰기 평가 인력 양성
- ✓ 글쓰기 대량 진단시스템 개발
- ✓ 글쓰기 자료 수집



# 1. '인공지능 활용 국어능력 진단체계 개발' 사업 소개

## 가. 사업 개요 및 진행 상황

### ■ 사업 개요

#### ✓ 글쓰기 자료 수집

- 논술문 자동 채점 시스템 모델 개발을 위해 활용되는 기존 데이터 규모 이상 확보 필요 (예) 영어권 ASAP 데이터: 6-8개의 쓰기 과제에 대해 1만7천 건 이상 구축
- 2023년 5천 건 이상을 목표로 2025년까지 매년 단계적으로 증량
- 데이터 다양성 확보를 위해 다양한 수준의 참여자 선정 중요
- 공공기관의 협약 등을 통한 쓰기 평가 참여자의 안정적인 확보가 중요



# 1. '인공지능 활용 국어능력 진단체계 개발' 사업 소개

## 가. 사업 개요 및 진행 상황

### ■ 사업 개요

#### ✓ 글쓰기 평가 인력 양성

- 채점 과정 체계화와 채점자 교육 방안
- 채점의 절차와 유의사항을 상세화한 채점 안내서 개발이 중요
- 채점 신뢰도 확보를 위해 채점 역량 평가를 통한 인증 방식으로 운영할 필요 있음.



# 1. '인공지능 활용 국어능력 진단체계 개발' 사업 소개

## 가. 사업 개요 및 진행 상황

### ■ 진행 상황

- 2022년 8월 11일 인천대학교와 국립국어원 인공지능 활용 국어능력 진단체계 개발 협력을 위한 업무협약 체결
- 2022년 12월 6일 '인공지능 활용 국어능력 진단체계 개발 협력' 협약 대학 연수회 개최
  - ✓ 참석자: 국립국어원 (원장, 담당 연구과장 및 연구원 7명), 연구진 (연구책임자 및 연구원 4명), 9개 협약 대학 글쓰기 담당 기관장과 실무 담당자 약 18명
  - ✓ 9개 협약 대학: 경상국립대, 부산대, 서울대, 인천대, 전남대, 전북대, 제주대, 충남대, 충북대



# 1. '인공지능 활용 국어능력 진단체계 개발' 사업 소개

## 가. 사업 개요 및 진행 상황

### ■ 추후 일정

- 2023년 1월: 협약 대학 협의체 구성(대표자, 실무자)
- 2023년 2월 중 1회: 협약 대학 대표자 및 실무자 협의회 개최
- 2023년 2월 중 1회: 사업 참여 교강사 온라인 설명회 개최



# 1. '인공지능 활용 국어능력 진단체계 개발' 사업 소개

## 나. 국립국어원과 각 대학의 역할

### ■ 국립국어원의 역할

- 글쓰기 진단지표 제공 (23년 2월 예정)
- 글쓰기 평가 인력 양성 과정 운영 (23년 4월부터 예정)  
→ 추후 논의 필요
- 학생 대상 홍보 동영상 제공 (23년 4월 예정)
- 글쓰기 자료 수집 관련 대학별 교강사 연수회 (23년 2, 8월 예정)
- 글쓰기 문제 제공 (23년 3월 예정)  
: 기초 연구에서 개발한 문제 10개 내외 제공 예정



# 1. '인공지능 활용 국어능력 진단체계 개발' 사업 소개

## 나. 국립국어원과 각 대학의 역할

### ■ 각 대학의 역할

- 글쓰기 평가 인력 양성 과정 수강 협조 (추후 논의 필요)  
→ 국립국어원 글쓰기 말뭉치 구축 사업 채점 인력으로 참여
- 글쓰기 자료 수집 협조
  - : 글쓰기 자료 수집 관련 대학별 교강사 연수회 참가
  - : 학생 대상 사업 안내 (사업 의의 및 개별 동의 필요성 등)
  - : 글쓰기 자료 수집



## 2. 국어능력 진단체계 개발 기초 연구 결과

### 가. 연구 개요 및 국내외 진단체계 운영 실태

#### ■ 연구 개요

##### • 연구 목적

인공지능을 활용하여 쓰기 능력을 진단하는 데 핵심적인 과제인 인공지능 학습용 쓰기 자료를 구축하고 진단체계 개발 및 운영 계획을 수립하기 위한 기초 연구



## 2. 국어능력 진단체계 개발 기초 연구 결과

### 가. 연구 개요 및 국내외 진단체계 운영 실태

#### ■ 연구 개요

- 세부 목표
  - ✓ 인공지능 활용 국어능력 진단체계 개발 및 운영을 위한 국내외 실태 분석
  - ✓ 인공지능 학습용 쓰기 자료 확보를 위한 원천 자료의 시범적 구축
  - ✓ 인공지능 활용 국어능력 진단체계 개발 및 운영 계획 수립



## 2. 국어능력 진단체계 개발 기초 연구 결과

### 가. 연구 개요 및 국내외 진단체계 운영 실태

#### ■ 국내외 진단체계 운영 실태

##### • 자국민 대상

(1) 국내: 국민의 국어능력 실태 조사, 서울대학교 신입생 글쓰기 평가 평가를 활용한 초중등학교 글쓰기 능력 신장 방안 연구  
국어능력인증시험, 한국실용글쓰기 검정

(2) 국외: NAEP, SAT, ACT, GMAT, GRE

##### • 외국인 대상

(1) 국내: 한국어능력시험

(2) 국외: FCE(케임브리지 영어시험), OWA(호주 성인대상 글쓰기 평가), TOFEL



## 2. 국어능력 진단체계 개발 기초 연구 결과

### 가. 연구 개요 및 국내외 진단체계 운영 실태

#### ■ 국내외 진단체계 운영 실태

- 대규모 논술 평가 사례\_평가 문항의 특성

- (1) 의사소통 목적

- 정보 전달: 4개 평가에서 사용됨.

- 설득/논증: 13개 평가에서 사용됨.

- 정서 표현: 3개 평가에서 사용됨.

- ✓ '정보 전달'과 '정서 표현'은 다른 유형과 함께 사용됨.



## 2. 국어능력 진단체계 개발 기초 연구 결과

### 가. 연구 개요 및 국내외 진단체계 운영 실태

#### ■ 국내외 진단체계 운영 실태

- 대규모 논술 평가 사례\_평가 문항의 특성

##### (2) 세부 논제 유형

묘사/설명: 4개 평가에서 사용됨.

의견 제시: 11개 평가에서 사용됨.

분석 평가: 5개 평가에서 사용됨.

✓ '묘사 설명'은 '의견 제시'와 함께 사용됨.



## 2. 국어능력 진단체계 개발 기초 연구 결과

### 가. 연구 개요 및 국내외 진단체계 운영 실태

#### ■ 국내외 진단체계 운영 실태

- 대규모 논술 평가 사례\_평가 문항의 특성

- (3) 자료 제시 여부

- 자료 미제시: 10개 평가에서 자료 미제시

- 자료 제시: 10개 평가에서 자료 미제시

- ✓ 6개 평가에서 자료 미제시와 자료 제시를 함께 사용함.



## 2. 국어능력 진단체계 개발 기초 연구 결과

### 가. 연구 개요 및 국내외 진단체계 운영 실태

#### ■ 국내외 진단체계 운영 실태

- 대규모 논술 평가 사례\_평가 문항의 특성

- (4) 작성 분량

- 500자 미만: 4개 평가에서 제시한 분량

- 500자 이상: 7개 평가에서 제시한 분량

- 제한 없음: 5개 평가에서 제시한 분량



## 2. 국어능력 진단체계 개발 기초 연구 결과

### 가. 연구 개요 및 국내외 진단체계 운영 실태

#### ■ 국내외 진단체계 운영 실태

- 대규모 논술 평가 사례\_평가 문항의 특성을 고려한 본 과업 평가 문항은 다음과 같음.
  - (1) 의사소통 목적: 설득/논증
  - (2) 세부 논제 유형: 의견 제시
  - (3) 자료 제시 여부: 자료 미제시
  - (4) 작성 분량: 500자 이상



## 2. 국어능력 진단체계 개발 기초 연구 결과

### 나. 문항 개발 및 평가 실행

#### ■ 문항 개발: 쓰기 문항 개발

##### 1. 평가 문항 원안 개발

- 평가 목적: 성인의 기초 쓰기 능력 측정
- 문항 유형: 논증적 쓰기 문항 개발  
: 쟁점 제시, 제시문 제시 유형 개발
- 인공지능 활용 고려: 원시자료 수집(300건)을 위한 평가 문항 개발(주제가 유사한 3개 문항)



## 2. 국어능력 진단체계 개발 기초 연구 결과

### 나. 문항 개발 및 평가 실행

#### ■ 문항 개발: 쓰기 문항 개발

##### 2. 전문가 검토

- 제시문 제시  
: 읽기 부담/난도 상승, 가치 중립적 자료 기술 문제, 표절 및 자료 재진술 문제가 있어 개발 시 난점이 예상됨.
- 쓰기 주제는 적절함.
- 쓰기 분량 원안대로(1000자) 진행
- 답안 작성 시간은 60분에서 90분으로 변경



## 2. 국어능력 진단체계 개발 기초 연구 결과

### 나. 문항 개발 및 평가 실행

#### ■ 문항 개발: 쓰기 문항 개발

##### 3. 예비 평가

- 1차 예비 평가
  - 쟁점 제시: 수험자 부담, 특정 방향으로 답안 작성 유도  
→ 쟁점 목록 삭제, 발문 진술 방식 변경
  - 답안 작성 분량: 분량 부족 문제 제기  
→ 문항 유형을 변경하고 분량은 현행 유지
- 2차 예비 평가
  - 답안 작성 분량: 분량 부족 문제 없음.
  - 평가 문항의 적절성: 적절함.



## 2. 국어능력 진단체계 개발 기초 연구 결과

### 나. 문항 개발 및 평가 실행

#### ■ 문항 개발: 쓰기 문항 개발

##### 4. 본 평가

- 4차 모집에 걸쳐 글쓰기 평가 답안 300건 확보  
(글쓰기 자료 수집에 어려움 있음.)
- 안정적으로 자료 수집이 가능한 온라인 플랫폼 구축



## 2. 국어능력 진단체계 개발 기초 연구 결과

### 나. 문항 개발 및 평가 실행

#### ■ 문항 개발: 평가 준거 개발

- 평가 지표  
: 내용 / 조직 / 표현
- 평가 준거  
: 문제 상황 이해, 주장, 근거의 적절성, 근거의 충분성, 다른 입장에 대한 고려 / 글 전체 구성, 문단 구성 요건 / 문장과 어휘, 어문규범과 관습



## 2. 국어능력 진단체계 개발 기초 연구 결과

### 나. 문항 개발 및 평가 실행

#### ■ 평가 실행

##### 1. 쓰기 평가 실행

- 연구 참여자: 총301명(대학생, 직장인 / 남녀 비율 50%)
- 쓰기 환경: 온라인 쓰기 플랫폼 내 문서편집기 활용
- 문항 배정: 3개 문항을 임의로 배정
- 시간 관리: 90분 이내 작성 (문항 확인과 동시 타이머 작동함.)



## 2. 국어능력 진단체계 개발 기초 연구 결과

### 나. 문항 개발 및 평가 실행

#### ■ 평가 실행

##### 2. 전문가 채점 체계

- (1) 채점자 섭외
- (2) 채점자 훈련
- (3) 1차 채점: 공통 채점, 온라인 채팅방 운영을 통해 채점자 간 일치도 높임.
- (4) 결과 분석 및 피드백: 채점자별 채점 척도 사용 특성 보고
- (5) 2차 채점
- (6) 결과 분석: 동일 문항 채점자 간 편차 분석
- (7) 추가 채점: 표준편차가 벌어진 글을 제3의 채점자가 재채점



## 2. 국어능력 진단체계 개발 기초 연구 결과

### 나. 문항 개발 및 평가 실행

#### ■ 평가 실행

#### 3. 채점 결과 및 통계 분석

- 문항 내적 일관성 신뢰도: 양호  
→ 동형 문항 추가 개발 및 문항 동등화 연구 필요
- 채점자 간 신뢰도: 우수  
→ 채점 객관도 향상을 위한 지표 고도화
- 채점 준거 적합도: 평가 준거 1(문제 상황 이해), 5(다른 입장에 대한 고려)의 준거 적합도가 낮음.  
→ 지표별 가중치 부여를 위한 연구 필요



### 3. '글쓰기 자료' 수집 방안 논의

#### 가. '글쓰기 자료' 성격

- 글의 성격: 논증적 글쓰기
  - ✓ 국어원에서 10개 내외 문제 제공 예정
  - ✓ 학생들이 사전에 준비할 수 없도록 문제 보안 필요
  - ✓ 다양한 성격의 글이 수집되는 것이 좋음.
- 분량: 1,000자
- 시간: 90분 (과제의 경우 시간 제한 불가능)
  - ✓ 원천 자료의 시범적 구축에서는 온라인플랫폼을 활용하여 시간 관리가 가능하였음.
- '글쓰기 자료' 활용
  - ✓ 국어원에서 별도의 피드백 없음. → 교강사 재량으로 활용



### 3. '글쓰기 자료' 수집 방안 논의

#### 나. '글쓰기 자료' 수집 분량 및 기한

- 협약 대학교 글쓰기 자료 수집안
  - 수집 분량: 총 7,000건(목표)
    - ✓ 협약기관 대표자 회의(2월) 시, 학교별 수집 수량 조정 협의
  - 수집 시기: 교과 일정에 따른 학기별 수집
  - 수집 방법: 대학 자율 수집(예: 수업 중 글쓰기 수집, 과제(리포트) 수집, 대학별 글쓰기 행사 또는 대회로 수집)
  - 수집 매체: 파일 수집, 온라인 수집 시스템(추후 운영 예정)
    - ✓ 대학에서 종이로 수집된 쓰기 자료는 자체적으로 전산 파일로 변환한 후 제출



### 3. '글쓰기 자료' 수집 방안 논의

#### 나. '글쓰기 자료' 수집 분량 및 기한

- 인천대학교
  - 수집 분량 및 시기: 학기당 500건 목표
    - ✓ 학생의 개인정보 및 저작권 활용 동의 필요하므로 수집 분량은 조정될 수 있음.
  - 수집 방법: 글쓰기 강좌 내 과제로 수집
  - 제출 방법: 파일 수합하여 국어원에 제출



### 3. '글쓰기 자료' 수집 방안 논의

#### 다. '글쓰기 자료' 수집 가안

- 고려할 사항
  - ✓ 자료 수집 분량 및 성격
  - ✓ 강의 계획서 통일 여부
  - ✓ 논증적 글쓰기 강의 여부
- '글쓰기 자료' 수집 방안 유형
  - ✓ 전 강좌 vs. 일부 강좌
  - ✓ 동일 주차 시행 vs. 자유 주차 시행

### 3. '글쓰기 자료' 수집 방안 논의

#### 다. '글쓰기 자료' 수집 가안

1

2

- 전 강좌 시행: 글 제출에 동의하지 않는 학생 수를 고려하여, 글쓰기 전 강좌에서 시행

- 동일 주차에 시행  
: 강의계획서에 동일 주차 과제 안내  
: 논증적 글쓰기 이론 강의 후, 과제로 활용  
→ 글쓰기 문제 보안에 효율적

- 자유 주차에 시행  
: 강의계획서에 교강사별로 자유 주차 과제 안내  
: 논증적 글쓰기 이론 강의와 별개 과제로 활용  
→ 다양한 수준의 글 수집에 효율적

### 3. '글쓰기 자료' 수집 방안 논의

#### 다. '글쓰기 자료' 수집 가안

3



- 일부 강좌 시행: 학생들의 적극적인 참여를 독려하여, 일부 강좌에서만 시행



- 동일 주차에 시행  
: 강의계획서에 동일 주차 과제 안내  
: 논증적 글쓰기 이론 강의 후,  
과제로 활용  
→ 글쓰기 문제 보안에 효율적

4



- 일부 강좌 시행: 학생들의 적극적인 참여를 독려하여, 일부 강좌에서만 시행



- 자유 주차에 시행  
: 강의계획서에 교강사별로 자유 주차 과제 안내  
: 논증적 글쓰기 이론 강의와 별개 과제로 활용  
→ 다양한 수준의 글 수집에 효율적



### 3. '글쓰기 자료' 수집 방안 논의

#### 라. '글쓰기 자료' 수집 방안 논의

- 글쓰기 자료 수집 방안 중 선택 or 또 다른 방안 ?
- 확정 기한: 강의계획서 탑재 전
- '글쓰기 자료' 제출 기한: 매 학기 16주차까지

감사합니다.

---