

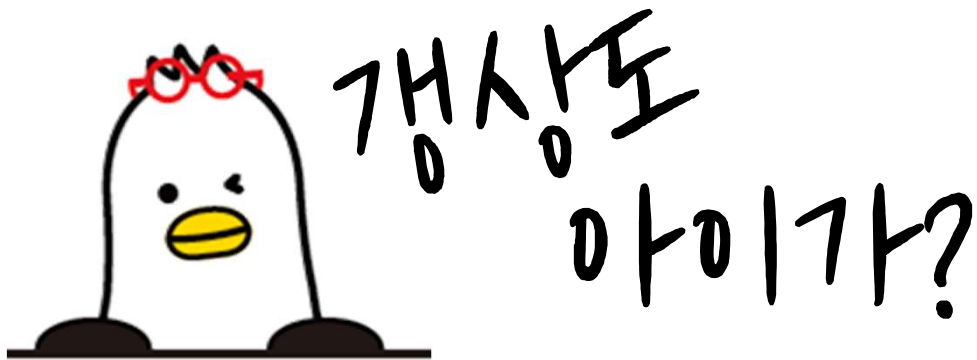
“경상도 아이가?”

경상도 사투리 발음 평가 서비스

프로젝트 기반 AI 빅데이터
K-digital training/부산대학교

3조 은서희, 장정원, 정대현









목차 소개

1. 시장 조사
2. 기획 의도
3. 최종 산출물
4. 로드맵
5. 개발 일정
6. 팀소개
7. 데이터
8. 주요 기술
9. Q&A

Z세대?

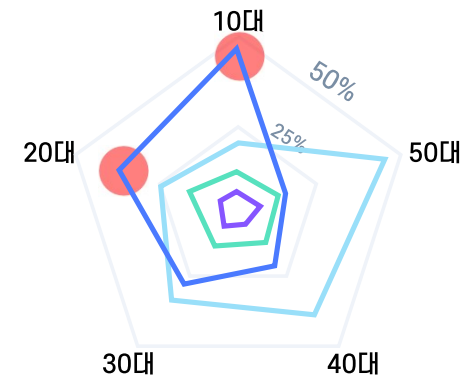
M세대를 뒤잇는 세대이며 일반적으로 1990년대 중반에서 2000년대 중반까지 출생한 세대
(한경 경제용어사전)

	1950	1960	1970	1980	1990	2000
세대 구분	베이비붐 세대	X세대	밀레니얼 세대(M세대)			Z세대
출생 연도	1950~1964년	1965~1979년	1980~1994년			1995년 이후
인구 비중	28.9%	24.5%	21%			15.9%
미디어 이용	아날로그 중심 	디지털 이주민 	디지털 유목민 			디지털 네이티브 
성향	전후 세대, 이념적	물질주의, 경쟁사회	세계화, 경험주의			현실주의, 윤리 중시

자료: 통계청·맥킨지코리아, 그래픽: 콘텐츠

주 여가 활동

- 휴식활동(TV, 산책 등)
- 취미, 오락활동
- 스마트폰 이용
- 문화예술관람



자료: 한국과학기술통신부(2021)

공통된 관심사를 기반으로 온라인에서 연결과 소속감을 느끼는 Z세대

공통된 정체성을 기반으로 온라인에서 뭉치는 Z세대

자신의 정체성을 유형화하는 '레이블링 게임'이 유행하고
이를 바탕으로 온라인상에서 익명의 만남과 소통 활발

나

- ESFP MBTI 성격유형검사
- 99학번 진학번테스트
- 만취한 장비 곧대레벨테스트
- 인트 캔디 SPTI 스넥 성격테스트



(이미지출처: 카카오톡 캡처)
자료 : 플레이디 (2021)



#연택트 소통
#성향 이해

서로를 잘 몰라도 함께 공유할 수 있는 연택트 소통과 성향 이해가 활발

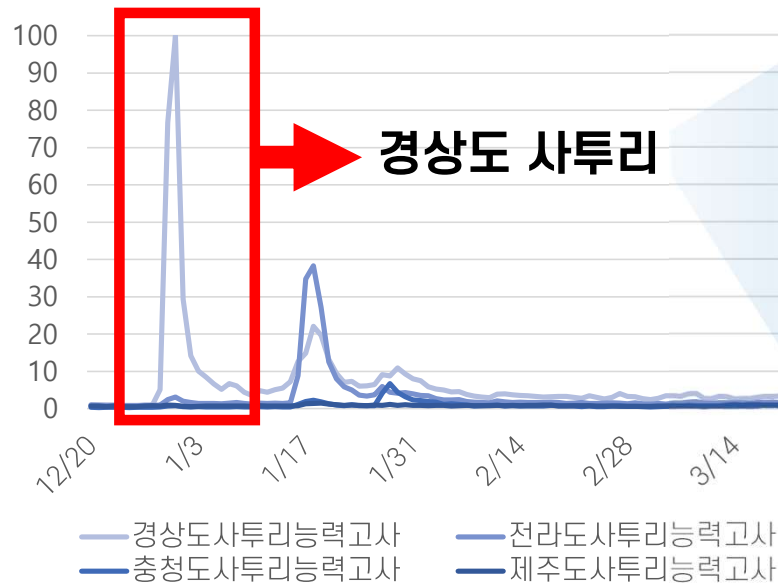
사투리는 Z세대 사이에서 지속적으로 각광받고 있는 콘텐츠



자료 : 구글 이미지 검색(2022)

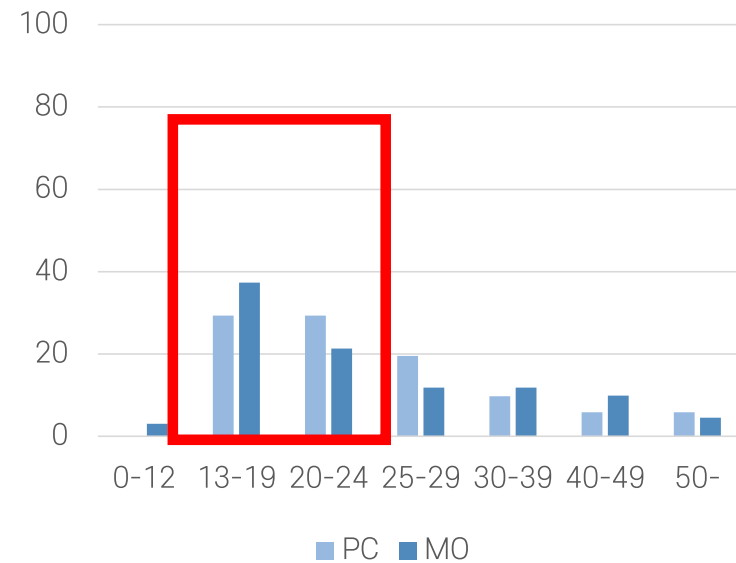
지역별 “사투리 능력고사” 키워드 검색 추이 확인 시,
“경상도 사투리”에 대한 관심도가 가장 높음

“사투리능력고사” 지역별 키워드 네이버 검색 추이



자료 : 네이버 데이터랩(2022)

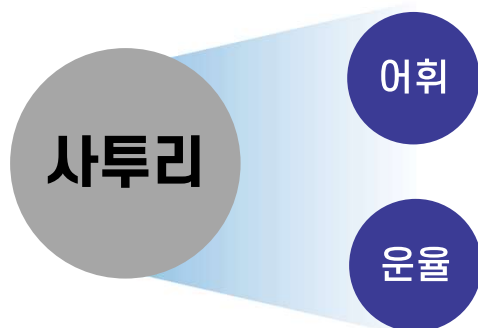
월간 검색수 사용자 통계 (최근일 기준) / 나이대(%)



자료 : 네이버 검색광고 “경상도사투리능력고사” 키워드 데이터

10 – 20대 Z세대 타겟으로 모바일용 웹서비스를 제작 계획

사투리는 크게 어휘와 운율로 구분되는데,
경상도 사투리는 성조 언어라는 큰 특징이 있음



사투리 요소

경상도 사투리 특징

경남 방언은 **성조언어**이다. (김차균, 1988)

“이 방언은 **음의 고저**에 대해서는 매우 민감하다.”
경남 방언의 음운론, p.40 (이근열, 1997)

①	가 가 가 가 ?	그 애가 그 애니?
②	가 가 가 가 가 ?	그 아이가 (성)이 '가'씨니?
③	가 가 가 가 가	'가'씨가 가서
④	가 가 가 가 가	그 애가 가져 간 다음에
⑤	가 ! 가 가 , 가 가 !	가래! 가서, 가져가렴!

경상도 사투리의 특징인 운율에 집중한 경상도 사투리 발음 능력 평가를 제작하기로

기존 사투리 퀴즈

경상도 사투리 능력 고사

성명 : 수험번호 202207200027

5. 다음 문장의 억양을 음계로 올바르게 표현한 것은?

어느정도높이까지올라가는거예요?

① 도레미파솔라시도레미파솔라시도

② 도레미파솔라시도레미파솔라시라

③ 도레미파솔파미레미파솔라시도레

④ 도레미파솔라시도시라솔라시도레

⑤ 도레미파솔라시도레미파솔파미레

① 도레미파솔라시도레미파솔라시도

② 도레미파솔라시도레미파솔라시라

③ 도레미파솔파미레미파솔라시도레

④ 도레미파솔라시도시라솔라시도레

⑤ 도레미파솔라시도레미파솔파미레



딥러닝 음성 분석을 통해
운율을 확인 할 수 있는
테스트를 만들자!

객관식 문제 형태로 경상도 사투리의 특징인
운율을 테스트하기 어려움

음성 분석을 할 수 있는 웹서비스를 통해 경상도 사투리의 운율을 직접 확인 가능

사투리 분석 검사

사투리를 사용하고 있는지
확인해보세요!



카카오 로그인

구글 로그인

버튼을 누른 후,
아래의 글을 따라 읽으시오.

1. 가가 가가?



버튼을 눌러
녹음을 시작해주세요.

성적표가 발급되었습니다.

재미로 하는 경상도 사투리 능력 고사에
응시하느라 고생 많았습니다.

친구에게 공유하고
성적표를 확인해 보세요!
1초도 안 걸려요!



카카오톡 트위터 페이스북 주소복사

테스트 다시하기

경상도 사투리 능력 고사 성적증명서

수험번호	성명
202207190022	***

등급	점수	백분위
9	2	100%

2022년 07월 19일

경상도사투리

재입 결과는 아랫부분에 있어요!

점수대별 하고싶은 친구에게 공유하기!



카카오톡 트위터 페이스북 주소복사

테스트 다시하기

문제별 정확도 결과



1.블루베리 스무디



경상도 마이가!!
정확도90%

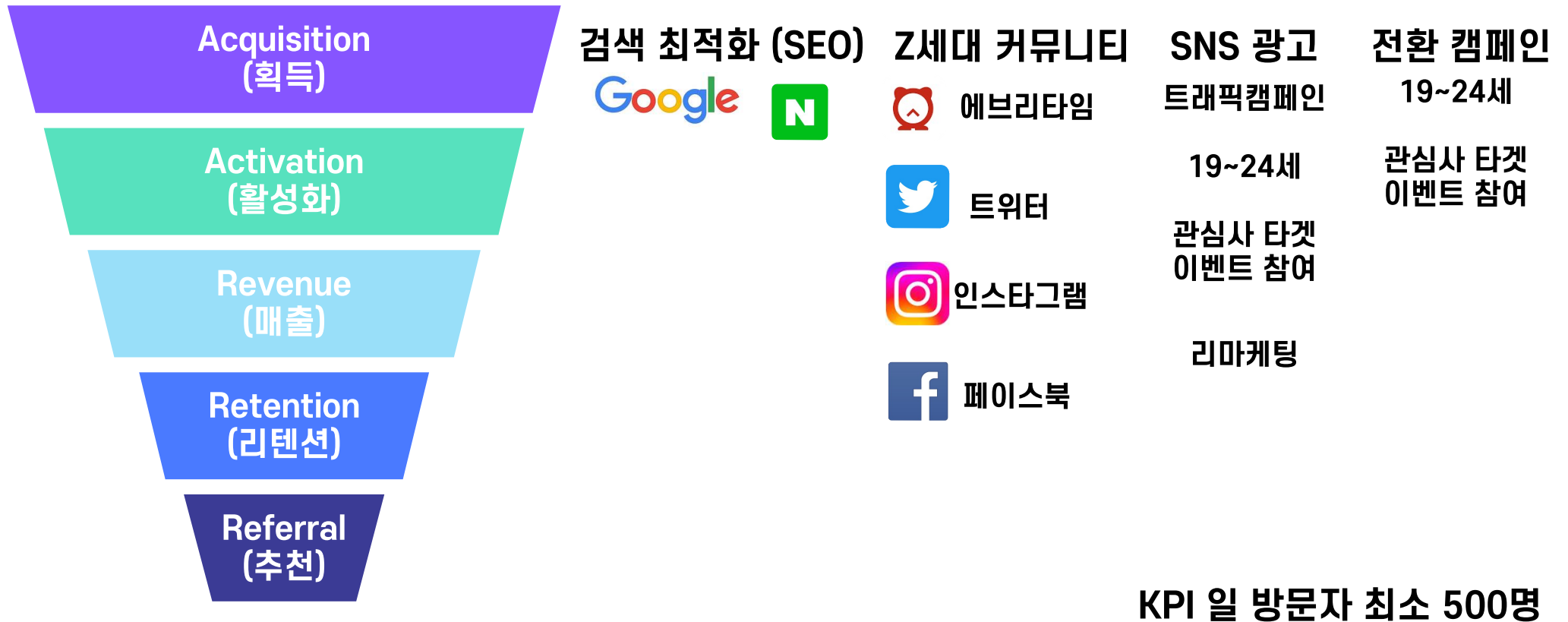
테스트할 문제 예시



1. 가가 가가? (그 사람이 아까 그 사람이었어?)
2. 2^e e^2 2^2 e^e
3. 블루베리 스무디 부산  vs 서울 
4. 어디까지 올라가는거예요?
5. 문디 가스나야.. 만다꼬 질질짜맨서 방 기팅이에 썰리 공가가 있노?
 니가 그카이 가가 그카지 니가 안그카른 가가 만다꼬 그칼끼고
 글마 끝베이가튼데 단디 정니하고 그른너른 재끼뿌고 퍼득 이자뿌라
 영 파이다 금마 아이라도 까리하고 혼빵가는 아덜 천지빠까리다
 고다꼬삭쌔미 추잡꾸로 그기머꼬 ? 글그치게스리 내한테 함 자피바라
 고마썰리마 어데 널짜뿌가꼬 궁덜 주차뵐라니깐—
 마 쫌 인자 고마질질짜라 앵가이 햇쓰니까네...

4 홍보 계획

Z세대들이 많이 사용하는 커뮤니티 활용 및 이벤트 참여 높은 Z세대 타겟으로한 SNS진행하여
초기 유입 확보할 것이며, 이후 자발적 공유를 통한 바이럴 마케팅 목표



장기적 사업 확장 가능성: 사투리 발음 교정 서비스



[현재 프로젝트]



경상도 사투리
발음 평가 서비스

제주도 사투리
발음 평가 서비스

전라도 사투리
발음 평가 서비스

충청도 사투리
발음 평가 서비스

사투리 발음 교정
서비스 출시

초기

중기

장기

데이터 수집 이후 지역을 확대하여 지역별 발음 능력 평가를 제작할 것이며, 사투리 억양 교정이 필요한 취업준비생들을 위한 서비스를 출시할 계획

서울신문

"취업위해 사투리 교정...500만원 썼습니다"[이슈픽]

입력 : 2021-03-27 19:25 | 수정 : 2021-03-27 20:05

◆취업준비생들도 사투리가 고민스럽다?

취직을 준비할 때도 사투리는 여전히 고민거리로 등장한다. 취업포털 커리어에 따르면 취업준비생 절반 이상(58.9%)이 사투리 교정은 필요하다고 응답했다. 사투리 교정이 필요한 이유에 대해 80%가 '표준어가 사회생활에 도움이 되기 때문'이라고 답했다. 그 다음으로 면접에서 자신감을 얻을 수 있어서 (15.7%), 사투리 때문에 느끼는 소외감이나 이질감을 없앨 수 있어서 순이었다. 더구나 응답자 본인이 기업 인사담당자라면 '구직자의 사투리가 채용여부에 영향을 줄 것 같다'라는 답이 56.5%에 이르렀다. 반면

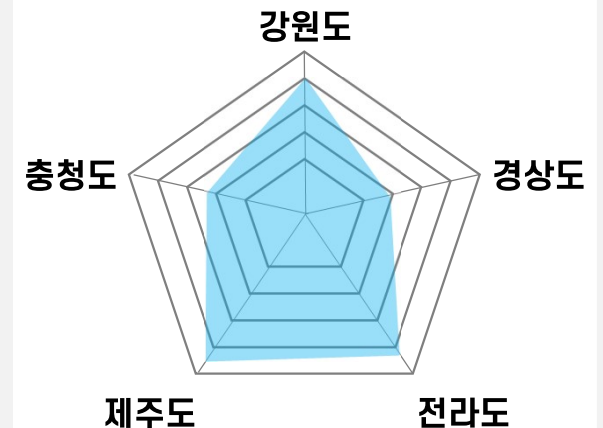
자료 : 특집부, "[다시,사투리] 설문조사로 본 사투리의 현실", 매일신문, 2021.01.15

사투리 교정 서비스 결과

표준어 정확도

35%

당신은 **제주도** 사람이군요!!



최종 프로젝트 개발 일정

구분	기간	활동	비고
1. 기획	7/18~7/21	기획의도 역할분담 정의	
2. 데이터 수집	7/22~7/24	데이터 다운로드(2TB)	
3. 모델 사전 학습	7/25~7/28	음성 모델 인식을 위한 CNN+BiGRU 학습	
4. 데이터 전 처리	7/28~7/31	방언 발화 데이터 전처리	
5. 모델링	7/28~8/5	방언 데이터 패턴 검출을 위한 딥러닝 모델 제작	중간점검
6. 모델 개선	8/6~8/19	모델 개선과 학습	
7. 웹서비스 구축	8/10~8/19	Django를 이용한 배포 웹서비스	
8. 발표 준비	8/10~8/19	PPT 제작 최종 발표 준비	
분석한 내용 PPT제작	7/18~8/25 총 5주		



장정원

PPT 기획안 구성
음성 데이터 수집
데이터 전처리



은서희

PPT 기획안 구성
음성 데이터 수집
음성 분석 모델 학습



정대현

서버 구축
학습 모델 구축
웹사이트 구현
(Django)

표준어와 경상도 사투리 데이터

데이터	상세 정보	데이터 구축년도/ 데이터 구축량	비고
한국어 방언 발화(경상도)	https://aihub.or.kr/aihubdata/data/view.do?currMenu=115&topMenu=100&aihubDataSe=realm&dataSetSn=119	2020년/432만	
숫자가 포함된 패턴 발화 데이터	https://aihub.or.kr/aihubdata/data/view.do?currMenu=116&topMenu=100&aihubDataSe=ty&dataSetSn=484	2021년/음성데이터: 13,342시간	수도권&경상도
화자 인식용 음성	https://aihub.or.kr/aihubdata/data/view.do?currMenu=115&topMenu=100&aihubDataSe=realm&dataSetSn=537		수도권&경상도
방송 콘텐츠 대화체 음성인식	https://aihub.or.kr/aihubdata/data/view.do?currMenu=115&topMenu=100&aihubDataSe=realm&dataSetSn=463		라디오만, 진행자 &environment: '배경음악' 제외

사용 예정 주요 기술



QnA

관광 이벤트 캠페인으로 제작하여
해당 지역 사투리 직접 체험하는 캠페인으로도 이용 가능



경상도 사투리 발음 평가 결과

95점

당신을 쯔 갯상도

사람으로

임명합니다



캠페인 프로세스

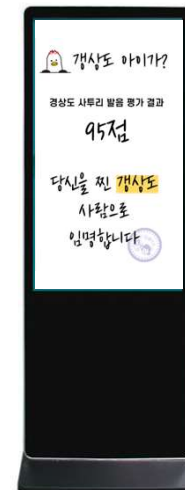


관광객

사투리
발음 테스트



점수에 따른
보상 이벤트

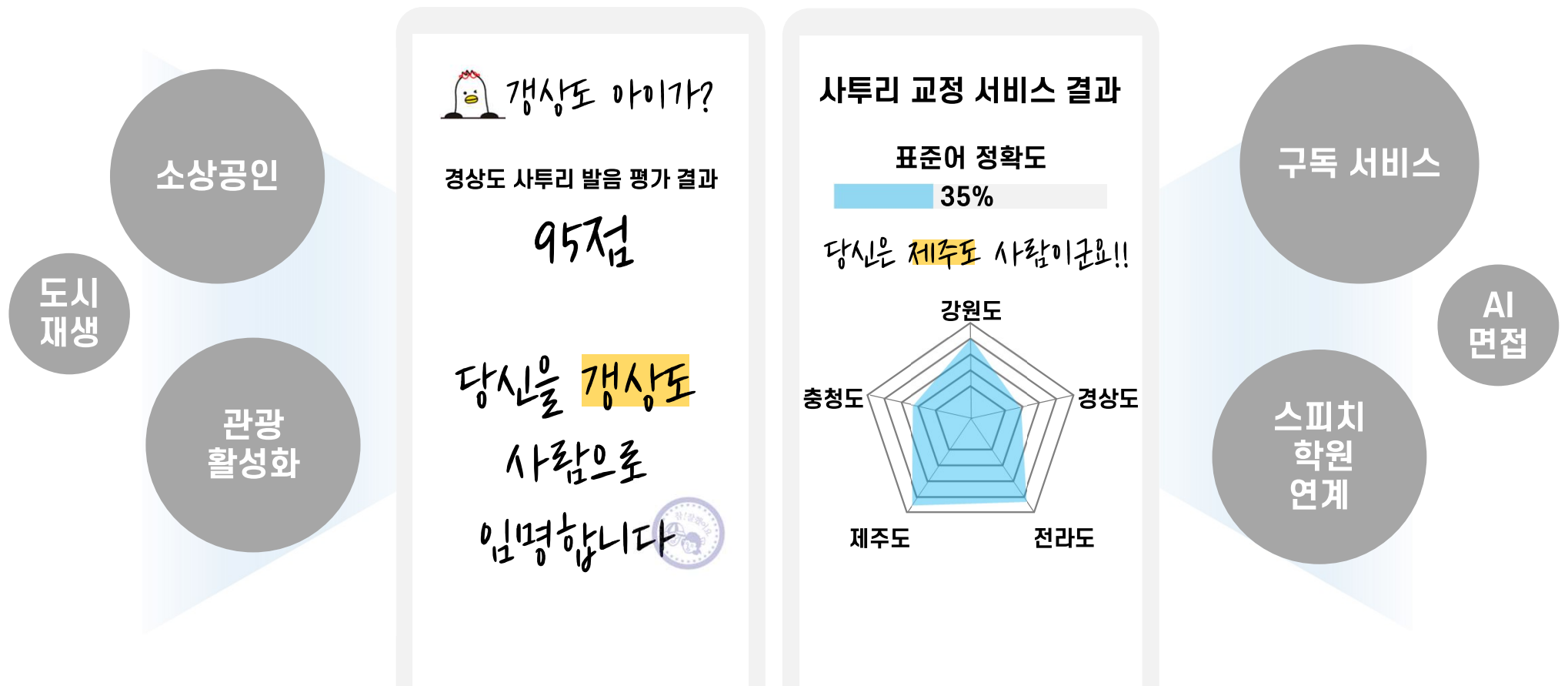


키오스크

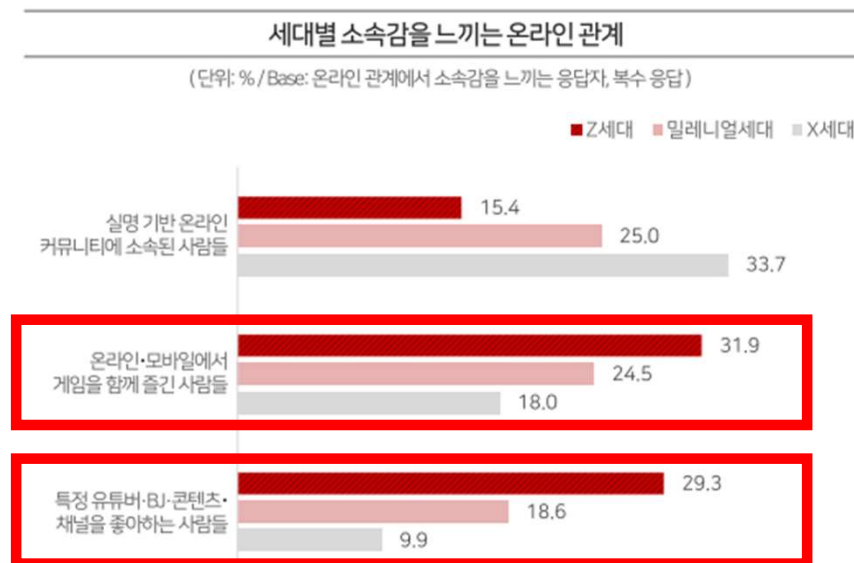
소상공인
연계 할인

지역 화폐

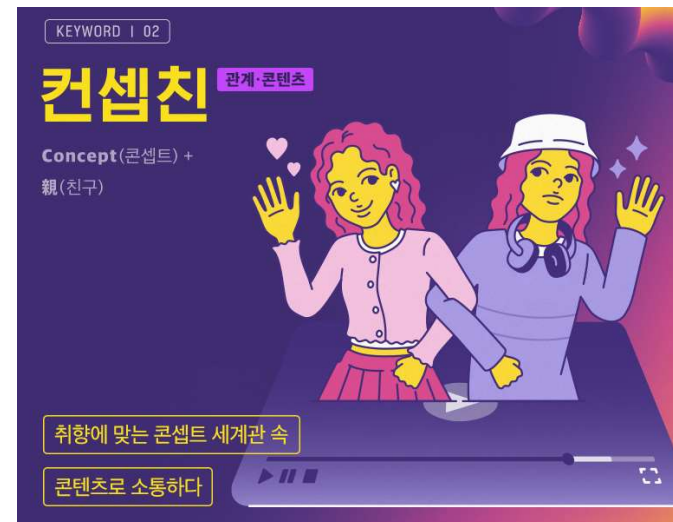
데이터 수집 후 지역별로 서비스를 제공하는 것에 그치는 것이 아니라,
데이터를 기반의 다른 서비스도 추가 제작하여 사회 가치 증진



Z세대는 깊은 오프라인 관계보다는 공통의 관심사를 기반으로 가벼운 온라인 만남 선호



자료 : 플레이디 (2021)



Z세대 성격과 성향으로 관계맺기 선호(18%)

자료 : 대학내일 (2022)

깊은 관계보다는 온라인 상에서 만난 익명의 사람과의 가벼운 상호작용을 선호하여
온라인 커뮤니티와 SNS상 비대면 커뮤니케이션 활발