**[성남행사 AI파인더]**

| **작성자/팀명:** | **크루** |
| --- | --- |
| **참가 부문:** | **학생부** |
| **소속 (참가 부문 학생부인 경우에만 작성)** | **판교고등학교** |

1. **프로젝트 정의**

| 1. **제안 배경**   성남시는 매년 다양한 문화행사를 개최하고 있으나, 시민들의 입장에서는 어떤 행사가 언제 열리는지 일일이 확인하기 어렵다. 또한, 행사가 자신의 관심분야인지 여부를 알기 위해서는 시간과 노력을 들여 사이트들을 뒤져야 한다. 현재 성남시 공식 블로그가 운영 중이지만, 사용자가 원하는 방식으로 AI시대에 걸맞게 서비스가 제공되고 있다고 보기 어렵다. 시민들은 “오늘 어떤 행사에 참여하는 것이 좋을까?”란 질문에 “지금 바로 답을 얻고 싶다”는 욕구가 크지만, 막상 행사 주최별(시청, 도서관, 유스센터, 성남시문화재단 등의 홈페이지와 보도자료)로 흩어져 있어서 정보를 찾는데 많은 시간이 소요된다. 시민들이 관심 있는 행사를 제때 알지 못하면 참여할 기회를 상실하게 된다. 시민들을 위한 행사임에도 불구하고 참여율이 낮다면 아무리 좋은 행사도 실패작이 되고, 시민들은 좋은 행사를 놓쳐버리는 안타까운 상황이 연출된다.  이에 행사 주최자와 수요자 간 매칭을 돕기 위해 개인별로 관심있는 행사를 정확도 높게 추천해 주는 ‘성남행사 AI파인더’ 앱을 기획하였다. 앱의 기본원리는 사용자의 온라인 활동내역(관심 행사 조회 수 등)을 입력받아 점수화 하고, LLM을 통해 선호할 행사를 추천하는 것이다. 문화행사에 관심이 많은 고등학생의 사례를 보자. “오전 10시부터 오후2시까지는 학원일정이 없는데, 대중교통으로 이동 가능한 범위 내의 환경과 관련된 행사 없을까?” 궁금해 한다. 단순한 알고리즘으로 만든 앱은 이러한 질문에 효과적으로 답할 수 없으며, LLM을 활용해야 한다. 기획한 ‘성남행사 AI파인더’를 성남시가 도입한다면, 사용자(시민/관광객)가 행사정보를 “검색 후 다운로드 확인”하는 성가신 과정을 거칠 필요가 없다. AI챗봇 질문 한 마디에 사용자가 원하는 바를 즉시 확인할 수 있다. 기관의 사이트를 크롤링하여 실시간 정보를 제공할 수 있기 때문이다. ‘성남행사 AI파인더’ 앱을 통해 성남시에서 열리는 다양한 행사에 대한 정보가 시민 개인별로 맞춤형으로 제공된다면 문화복지가 월등히 제고될 수 있을 것이다.   1. **프로젝트 개요**    * **차별성 및 필요성** : 성남시는 ‘성남시 공식 블로그’를 통해 다양한 성남시 주최 문화 행사 데이터를 공유하고 있지만 현재 정보 제공은 검색 중심의 정적 웹페이지에 한정되어 있어 일반 시민이나 관광객의 자발적 접근이 어렵다. 본 프로젝트는 기존 콘텐츠를 대화형 인터페이스로 재구성하고, 대규모 언어모델을 활용해 사용자 질문에 따라 유연하게 설명할 수 있는 성남시 주관의 행사 특화 챗봇을 구축할 예정이다. 이는 단순한 정보 나열을 넘어서 사용자 관심 기반 큐레이션, 연계 콘텐츠 추천 등 능동적인 전달 방식을 제공함으로써 기존 정보 시스템과 명확히 차별화된다.   1) 단일 지자체 특화 : 성남시 전용 데이터만 통합하여 답변 정확도 ↑  2) 데이터-지식 융합 : 구조화 데이터 + 시 보도 자료 + FAQ(Frequently Asked Question) + 성남시청 웹사이트의 이벤트 설명 페이지  3) 비전문가 UX: 웹 챗(CHAT)-음성(시니어;어르신) 지원, PWA(Progressive Web App) 기반 다운로드 가능한 앱 지원   * + **핵심내용 및 특징** :  1. 성남시, 성남시 도서관, 성남시 유스센터, 성남시청소년청년재단, 성남시 문화재단 등 성남시 산하 기관의 행사 데이터를 한 번에 처리할 수 있음 2. 행사 장소 등 위치 데이터를 지도 위 시각화 3. 한국어/영어(외국인), 초등학생용 쉬운말 모드 지원 4. 행사 시간, 장소, 참여대상, 접수기간 등 정보를 통합적으로 처리해 한 번의 대화 안에 모든 조건을 고려할 수 있음 5. 추천 알고리즘을 활용해 실시간으로 사이트를 크롤링해 사용자가 좋아할 만한 이벤트를 폰 알림으로 추천할 수 있음 6. 사용자가 선호하는 행사 시간, 장소, 행사 주제 등의 데이터를 수집해 성남시에 제언 |
| --- |

1. **기술 구현 계획**

| 1. **AI분야**   - LLM 기반 챗봇 : GPT 등 대규모 언어 모델(LLM)을 활용하여 성남시의 예정된 행사 정보에 대해 자연어 질의응답 및 해설 제공할 수 있다. 사용자가 “오전 10시부터 오후2시까지는 학원일정이 없는데, 대중교통으로 이동 가능한 범위 내의 환경과 관련된 행사 없을까?”, “이번 주말에 가족끼리 갈만한 행사 추천해줘”와 같이 자유로운 형식으로 질문하면 챗봇은 그 의도을 파악하고 요약, 설명해주는 기능 제공한다.  - 사용자 맞춤형 정보 추천 : 사용자의 연령, 관심사, 일정, 위치 등 입력된 정보를 바탕으로 개인별로 가장 적합하고 흥미를 가질만한 행사 정보를 선별하여 추천한다.  - 크롤링 정보 선별 및 정제 : 성남시청, 도서관, 유스센터 등 다양한 기관의 웹사이트에서 자동 크롤링한 데이터 중 중복되거나 불필요한 정보를 AI가 지능적으로 필터링 및 정제하여 신뢰도 높은 정보만 사용자에게 제공한다.   1. **프로토타입 설계도**  * 사용자 입력: 자연어로 사용자가 질문, 추천 입력 * LLM: GPT 등 언어 모델을 활용하여 사용자의 질문을 이해하고 맥락에 맞는 답변 생성, 선별된 데이터를 통해 사용자 설정(연령,관심사 위치 등)에 적합한 행사 추천 * DATA: 파이썬을 통해 성남시 행사 주최별 사이트를 크롤링해 정보 수집, AI를 이용하여 중복되거나 불필요한 정보 선별, 정제 * 외부 데이터 수집: 성남시청, 도서관, 유스센터 등 공식 웹사이트 크롤링, 정기적 데이터 수집 및 업데이트 * AI 기반 정보 선별 및 정제: 크롤링된 데이터의 중복 제거 및 품질 관리, 관련성 높은 정보만 선별하여 저장 * 응답: 크롤링된 데이터 내 텍스트 정보 가공하여 사용자에게 제공, 이미지 및 이미지 내 텍스트 인식해서 정보, 관련 링크(데이터 출처, 신청 링크 등) 제공 |
| --- |
| **3. 주요 기술 및 기능 설명**   * 데이터 수집 및 통합 모듈   + 웹 크롤링 : crawl4ai 모듈을 이용하여 웹사이트를 크롤링해 인공지능 맞춤형 데이터 수집. 파이썬 request 라이브러리를 활용해 서버 코드를 작성하고 라즈베리파이4에 서버 코드를 배포해 직접 크롤링 서버를 개발. 또는 따로 타사이트에서 제공하는 파이썬 서버 스크립트 배포 서비스 활용   + 데이터 정제 및 저장 : AI 기반의 정보 선별 및 정제 기능을 통해 중복되거나 불필요한 정보 필터링하고 정보 품질 관리 * LLM: OpenAI의 ChatGPT4 API 활용, Axios 모듈을 도입해 API 활용, 사용자 맞춤형 정보 추천: 사용자의 입력(연령, 위치, 관심사 등)을 기반으로 GPT4Rec 등의 모듈을 활용해 행사 정보를 선별하여 추천함 * 웹 개발: Vue 프레임워크 기반 Nuxt 메타프레임워크를 활용해 체계화된 웹 서비스 및 Vuetify 디자인 라이브러리를 활용해 일관적인 웹 페이지 디자인 제공 * 호스팅: Firebase 호스팅 서비스를 활용해 사이트를 제공, PWA(Progressive Web App)을 활용해 구글 플레이스토어와 에플 스토어에 등록 |

1. **프로젝트 계획안**

| 1. **프로젝트의 개요 및 목적**   성남시 산하기관(시청, 도서관, 유스센터 등)의 행사 정보를 실시간으로 수집 및 정제하고, LLM + RAG 기반의 인공지능 챗봇을 통해 개인화, 대화형 추천을 제공하는 웹서비스의 구축.   1. **예상 결과물**   ‘성남행사 AI파인더’ 챗봇 (웹사이트/앱), 오픈소스 ETL, 행사 관련 정보 통계 및 지도 - 시각화 모듈, 음성 모드 등, 접근성 향상 서비스, 시연용 발표자료, API 및 시스템 문서   1. **세부일정**  | 주차 | 마일 스톤 | 메인 작업 | 예상 산출물 | | --- | --- | --- | --- | | W1 | 기술 셋 확정 | - 크롤링 대상 사이트 확정 (robots.txt 점검)  - 프로젝트 아키텍처 구상 |  | | W2 | 크롤러 PoC & 데이터 스냅샷, 데이터 정형화 | - Scrapy/Playwright 또는 Crawl4AI 기반 크롤러 완성  - 크롤링된 정보를 HTML -> JSON으로 저장  - 크롤링 로그, 라이트레이트 정책 적용  - 필드 정규화 (날짜,장소,카테고리), 이미지 저장 | PoC 크롤러 코드, 서버, JSON기반 API, 정형화 된 DB | | W3 | 앱 개발 | - 디자인 초안 완성  - 챗봇 디자인 및 웹사이트/앱으로 배포 | 웹사이트 개발, 디자인 | | W4 | 챗봇 프롬프트 작성 | - 프로그램에 적합한 챗봇 프롬프트 작성  - API를 클라이언트 사이드와 연결  - Crawl4AI를 활용해 인공지능에게 크롤링한 데이터를 제공 | 챗봇 최초 버전 | | W5 | 시연 및 데모 버전 완성, 리즈,  문서화 | - 베타 테스터를 대상으로 만족도 조사 후 그 결과에 따른 프롬프트 조정  - 발표자료  - Firebase 배포 | 발표자료 및 데모 영상,  최종 데모 URL 및 APK | |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |