

# 지산학캡스톤디자인 프로젝트 수행계획서

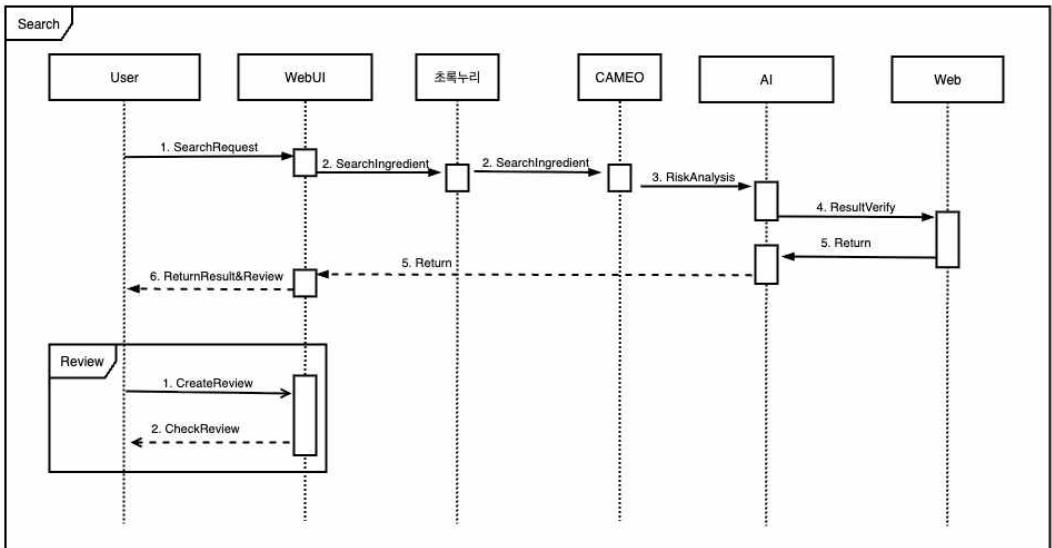
학생 팀별 작성용

프로젝트 수행팀 현황					
수행 학기	■ 2025년 9월~2025년 12월				
프로젝트명	■ 화학제품 혼합 안전도 검증 서비스				
팀명	■ Nemo				
교과목명	■ 지산학캡스톤디자인				
	학과	학번	성명	연락처	E-mail
팀장	정보통신공학과	2023111976	박수빈	010-9011-5676	qkrtnqls5676@gmail.com
팀원	정보통신공학과	2023112000	이연우	010-9873-3805	cg10046@naver.com
	정보통신공학과	2023111961	최형원	010-2711-8756	chrisychoe@dgu.ac.kr
	멀티미디어공학과	2022112456	김민영	010-8007-2964	a1637528@naver.com
지도교수	소속	SW교육원		넥스브이	
	성명	이길섭 교수		송환구 교수	
멘토	소속	(주)넥스브이		송환구 연구소장	

프로젝트	
프로젝트 개요	본 프로젝트는 생활 화학제품을 섞어 사용하기 전에 안전성을 신뢰 가능한 근거에 기반하여 즉시 판단할 수 있도록 돋는 웹 기반 "혼합 안전도 판정 서비스"를 개발하는 것을 목표로 한다. 가정에서 빈번히 사용되는 세제, 소독제, 식초 등은 조합에 따라 유해 가스를 발생시킬 수 있으며, 특히 차아염소산나트륨(일명 락스)은 산성 물질과 반응하면 염소가스, 암모니아계와는 클로라민을 발생시켜 화학성 폐렴이나 폐부종과 같은 급성 건강 피해를 초래할 수 있다. 그러나 현재의 정보 유통 구조는 제조사 고객센터, 지식 커뮤니티, 공공 캠페인 등으로 흩어져 있어 소비자가 혼합 직전에 신뢰할 수 있는 결론에 도달하기 어렵다. 이에 본 프로젝트는 사용자가 텍스트, 제품 라벨 사진(OCR), 음성 등으로 재료를 입력하면 공신력 있는 기관의 데이터만을 조회하여 혼합의 위험도를 금지, 주의, 불명확, 안전의 네 등급으로 판정하고, 근거 출처와 대체 사용법을 함께 제시하는 서비스를 설계·구현한다. 최종적으로 작동 가능한 프로토타입을 완성하여 정확도와 응답 속도, 사용자 이해도를 검증하고자 한다.
추진 배경 (자료조사)	1. 개발 배경 및 필요성

<p><b>( 및 요구분석)</b></p>	<p>비대면 생활과 위생 관리의 강화로 가정 내 소독·세정 행위가 증가하면서, 제품 혼합 사용에 대한 질문이 온라인 상에서 반복되고 있다. 일상에서 사용하는 세정제·소독제·산성 세제·식초 등 생활 화학제품은 조합에 따라 예기치 못한 반응을 일으킬 수 있다. 대표적으로 차아염소산나트륨(가정용 표백제 성분)은 산성 물질과 만나면 염소가스, 암모니아계 물질과는 클로라민을 생성해 급성 호흡기 손상을 유발할 수 있다. 이러한 위험 정보는 의학 저널과 MSDS에 명확히 정리되어 있으나, 실제 소비자는 "A 제품과 B 제품을 같이 써도 되는가?"라는 제품명 중심의 질문을 갖는다.</p> <p>현재 정보 환경에서는 사용자가 각 제품의 성분표를 직접 찾아 비교해야 하는데, 포장지 라벨의 가변 표기·브랜드별 레시피 차이·전문 용어(예: CAS, GHS 분류) 등으로 인해 즉시적인 자기 판단이 사실상 불가능하다. 그 결과, 커뮤니티 경험담이나 제조사 고객센터 답변에 의존하게 되어 근거의 일관성 부족과 오해가 반복된다.</p> <p>본 프로젝트는 이 정보 비대칭을 해소하기 위해, 사용자에게 제품명만 입력하게 하고 시스템이 공공 데이터(초록누리)에서 정확한 성분을 자동 추출한 뒤, 공신력 있는 반응성 데이터(CAMEO)를 교차 조회하여 표준화된 등급 판정과 행동 지침을 제공하는 것을 목표로 한다. 즉, 위험 지식은 이미 존재하지만 접근성과 해석성이 낮다는 현실적 문제를, 데이터 파이프라인과 자동 판정으로 해결한다.</p> <p><b>2. 선행기술 및 사례 분석</b></p> <p>데이터 소스 측면에서, 안전보건공단의 MSDS는 물질별 반응성·비호환성·응급조치·GHS 분류를 표준 서식으로 제공한다. NOAA의 CAMEO Chemicals는 물질 간 반응성 매트릭스와 대응 개요를 제공하여 성분 조합 위험 평가에 적합하다. 화학물질안전원(NICS)은 국내 규제·사고 대응 정보를 제공하고, IARC Monographs는 발암성 장기 위해 라벨링의 근거가 된다. 국내 초록누리는 전성분 공개 제품 DB로, 제품명→성분 매핑의 신뢰성이 있는 출발점이다. 그러나 이들 소스는 플랫폼이 분산되어 있고, 일반 소비자에게는 질의 양식이 물질 중심이어서 제품명에서 시작하는 질문에 바로 답을 주지 못한다.</p> <p>기존 서비스/관행 측면에서, 제조사 Q&amp;A나 커뮤니티(지식형 플랫폼)의 답변은 접근성이 높지만 근거 표준화·출처 일관성이 부족하고, 동일 질문이 반복되기 쉽다. 공공기관의 안전 캠페인은 경각심을 높이지만 사용 순간의 개별 조합에 대해 즉답형 도구를 제공하지는 못한다. 해외의 일반 검색·문답형 챗봇은 생성형 모델에 의존하여 환각(hallucination) 위험이 있고, 링크 출처가 신뢰 도메인과 혼재되는 문제가 있다.</p> <p>본 프로젝트의 차별점은 (1) 제품명→성분 정규화(초록누리), (2) 성분 간 반응성 자동 조회(CAMEO 등 화이트리스트 도메인), (3) 룰엔진 기반 등급화(금지·주의·불명확·안전) + 정형 템플릿 출력, (4) LLM은 요약·표현에 한정하고 결론 전 자가 서치(구글/빙) 결과 중 공신력 도메인만 추천 링크로 첨부하여 신뢰를 보강한다는 점이다.</p> <p>또한 초기에는 초록누리에 등록된 제품을 우선 지원하여 정확도와 속도를 확보하고, 미수록 제품 보완은 후속 단계로 이연해 서비스 가동성을 우선 확보한다. 특히·권리 측면에서는 공공 DB 링크 인용을 원칙으로 하고, 판정 로직과 정형 출력은 자체 설계로 차별적 요소를 확보한다.</p> <p><b>(참고 문헌)</b></p> <p>대한의사협회지(2023). 「가정용 표백제(차아염소산나트륨) 노출의 진단 및 치료」. 혼합 시 유해가스(염소·클로라민) 발생과 임상적 결과를 체계화.</p> <p>안전보건공단 MSDS. 물질별 반응성·비호환성·응급조치·GHS 분류를 표준 서식으로 제공; 룰엔진 규칙 정의의 1차 근거.</p> <p>NOAA CAMEO Chemicals. 물질 간 반응성·비호환성 및 대응 개요 데이터베이스; 성분군 매트릭스 일반화에 유용.</p> <p>IARC Monographs. 발암성 분류 체계; 단독 제품 분석 시 장기 위해 라벨링 근거.</p>
-----------------------------	---

	<p>화학물질안전원(NICS). 국내 규제·사고 대응 정보 및 검색 서비스; 국내 기준과의 정합성 확보.</p> <p>공공 안전 캠페인 자료(행정안전부 등). 생활화학제품 혼합 금지·환기·희석 등의 행동요령을 반복 고지; 사용자 인터페이스로의 전환 필요성을 시사.</p>
목표 및 내용	<p>본 프로젝트의 개발 목표는 첫째, 금지·주의 등 위험 사례에 대해 전문가 검수 50건 기준 근거 일치율 90% 이상을 달성하는 정확도를 확보하는 것, 둘째, 캐시 적중 상황에서 0.6초 내, p95 기준 1.5초 이내의 응답 성능을 달성하는 것, 셋째, 사용자 테스트에서 이해도와 신뢰도 평균 4.3/5 이상을 확보하는 것이다.</p> <p>최종 결과물은 반응형 웹 애플리케이션으로 구현되며, 필요 시 경량 관리자 콘솔과 PWA 기능을 제공한다. 사용자는 텍스트나 음성으로 재료를 입력하거나 제품 라벨을 촬영하여 업로드하면, 시스템이 식초를 아세트산으로 치환하는 등 성분 동의어를 정규화하고, 허용된 공공 도메인만을 검색하여 관련 반응과 독성 정보를 수집한다. 이어서 룰엔진이 금지, 주의, 불명확, 안전의 판정을 내리고, 정형 출력 템플릿이 이유와 근거 링크, 대체 절차를 서술한다.</p> <p>추가로 본 서비스의 1차 데이터 파이프라인은 다음과 같다:</p> <ol style="list-style-type: none"> <li>(1) 사용자가 텍스트를 입력하면 우선 초록누리 전성분 공개 제품 데이터에서 해당 제품의 성분표를 조회하여 가져온다(초록누리에 없는 상품 보완은 후속 단계로 이연).</li> <li>(2) 가져온 성분들을 내부 검색/규칙으로 서로 비교하여 위험 조합을 탐색한다(현재 사진에 제시된 서치 방식 적용).</li> <li>(3) 이 비교 결과와 축적된 근거만을 바탕으로 AI가 최종 판정을 내리되, 제공 데이터 외로 확대 추론하지 않도록 페르소나·출처 제한을 적용한다.</li> </ol> <p>서비스의 행태는 사용자와 시스템의 상호작용 흐름(검색-비교군 선택 또는 추가-결과)과, 로그인 후 마이페이지·신고 기록으로 이어지는 구조로 표현되며, 제공된 블록도와 시퀀스 다이어그램을 본문에 삽입하여 입력, 내부 처리, 결과 간의 관계를 명확히 제시한다.</p> <pre> graph TD     User((User)) --&gt; 검색[검색]     검색 -- 필수 --&gt; 비교군[비교군]     비교군 -- 선택 --&gt; 추가[비교군 추가]     추가 -- 필수 --&gt; 결과[결과]      User --&gt; 로그인[로그인(일반/소셜)]     로그인 --&gt; 마이페이지[마이페이지]     마이페이지 --&gt; 검색기록[검색 기록]     마이페이지 --&gt; 신고후기[신고(후기)]     신고후기 --&gt; 신고기록[신고 기록]   </pre>



### 1. 대안 도출 및 구현 계획

여러 구현 대안을 비교한 결과, 생성형 모델 단독 사용은 개발이 빠르지만 환각과 근거 불명확 문제가 있어 채택하지 않으며, 전수 룰·DB 수작업은 정확도가 높으나 구축·유지 비용이 커 핵심 룰에 한하여 부분 채택한다. 최종적으로는 도메인 제한 검색과 LLM 함수 호출을 결합한 하이브리드 방식을 선택한다.

주요 기능은 성분 정규화 사전 구축, 라벨 OCR 파싱, 화이트리스트 기반 Programmable Search 또는 기관 API 연동, 등급 판정 룰과 정형 출력 템플릿의 설계로 구성되며, 데이터 구조는 성분, 제품, 규칙, 질의 로그의 네 축으로 단순하고 확장 가능하게 정의한다. 전체 알고리즘의 흐름은 입력 정규화, 제한 탐색, 근거 스니펫 수집, 룰 판정, LLM 요약, 결과 렌더링의 순서로 진행되며, 플로차트로 상세를 제시한다.

또한 환각 방지 및 근거 보강을 위해 LLM이 결론을 내리기 전에 Google/Bing 자가 서치를 수행하되, 화이트리스트(공공·전문 기관) 도메인 결과만 필터링하여 추천 검색 링크로 결과 화면에 첨부한다. 관련 링크가 없으면 내부 데이터만으로 판단하고, 서비스 내 실제 사용자 후기를 같이 제시하여 AI 결과의 신뢰도를 높인다.

### 2. 설계의 현실적 제한요소(제약조건)

서비스는 공공 API 쿼터와 LLM 호출 비용의 제약을 받으므로 캐시와 배치 수집, 요약 저장으로 비용을 관리한다. 동작환경은 모바일 우선의 반응형 웹으로 정의하고, 저사양 기기에서도 목표 응답 성능을 만족하도록 설계한다.

개발환경은 단일 클라우드 호스트 위 Docker 컨테이너 스택(Nginx, Spring Boot, PostgreSQL, Redis)으로 구성하는 형태를 기본으로 하며, 개인정보는 최소 수집·암호화 원칙을 따르고, 출처 표기와 오정보 방지, 의료적 진단·처치 정보의 배제를 통해 법적·윤리적 요구를 충족한다.

### 3. 개발 환경

초기 파일럿은 단일 클라우드 호스트(2 vCPU / 4GB RAM) 위에 Docker 컨테이너로 배

	<p>포한다. 구성요소(Nginx, Spring Boot, MySQL, Redis)는 각각 별도 컨테이너로 분리하고, 배포·운영은 Docker Compose로 관리한다(트래픽 증가 시 Kubernetes 이행). 네트워크는 Nginx가 HTTPS 종단(SSL/TLS 인증서, 예: Let's Encrypt 자동 갱신) 후 내부 컨테이너로 프록시한다. 컨테이너 이미지는 멀티 스테이지 빌드(Jib/Buildpacks 선택)로 경량화하며, 헬스체크(/actuator/health), 리소스 제한(CPU/메모리), 재시작 정책을 적용한다.</p> <p>로그는 구조화(JSON)로 표준화하고, 비밀정보는 환경변수/Secret Manager로 분리한다. CI/CD는 GitHub Actions 등으로 빌드 → 테스트 → 이미지 푸시(Private Registry) → Compose 배포 파이프라인을 구성한다. 프론트엔드는 React 기반(또는 동등 웹 스택), 백엔드는 Spring Boot(Java/Kotlin, MVC 또는 WebFlux)로 구현한다. OCR은 Tesseract/클라우드 비전을 사용한다.</p>		
기대효과	경제적 측면에서 별도의 하드웨어 없이 소프트웨어만으로 도입이 가능하므로 초기·유지 비용을 절감할 수 있으며, 공공 안전 콘텐츠의 도달성과 활용도를 높여 제조사와 기관의 반복 상담 비용을 줄일 수 있다. 사회적 측면에서 유독 가스 노출 위험을 낮추어 응급·소방 부담을 경감하고, 생활화학 안전 리터러시를 향상시켜 이용자의 즉시 의사결정을 돋는다. 표준화된 경고 문안과 대체 절차를 제공함으로써 서비스 품질 전반을 향상시킨다.		
추진일정	세부 일정은 총 8주로 계획한다. 1~2주차에는 요구 도출과 정보 구조 설계, 동의어 스키마의 초안을 완성하고, 3~6주차에는 입력에서 판정과 출력까지의 전 과정을 연결하는 구현과 캐시, 로그, 관리자 기능을 개발하며, 7~8주차에는 통합 테스트와 성능·정확도 투팅을 거쳐 컨테이너 이미지 최적화·보안 스캔·Compose 배포로 시연용 배포를 마무리한다. 구성원 역할은 팀장 위주 팀원 전체가 일정과 품질, 대외 커뮤니케이션을 총괄하고, 프론트엔드가 접근성을 고려한 입력·결과 화면을, 백엔드가 API와 제한 탐색, 캐시·로그를, 데이터/AI 담당이 동의어 사전과 판정 룰 및 템플릿을, QA가 테스트와 이슈 관리를 전담한다.		
성과 창출 계획	항목	세부내용	예상(달성)시기
	Github	<a href="https://github.com/CSID-DGU/2025-2-DES4015-Nemo-01">https://github.com/CSID-DGU/2025-2-DES4015-Nemo-01</a>	
	논문제재 및 참가	저널 또는 학회명 :	
	SW등록		
	특허출원		
	시제품 (스토어에 등록)		

[양식 DES-02-2] 지산학캡스톤디자인 프로젝트 수행계획서