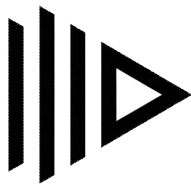


2. Analysis

Open Media Player



Open Media Player

Student No.	22313530
Name	BAE WON IL
E-Mail	baewonil@yu.ac.kr

[Revision history]

Revision date	Version #	Description	Author
04/06/2025	1.00	First Draft	BAE WON IL
02/05/2025	1.01	Live-Support 추가	BAE WON IL
07/05/2025	1.02	Minor Change (일자 추가, NFR 추가)	BAE WON IL
07/05/2025	1.03	Minor Change (Reference 수정)	BAE WON IL

= Contents =

1. [Introduction](#)
2. [Use case analysis](#)
3. [Domain analysis](#)
4. [User Interface prototype](#)
5. [Glossary](#)
6. [References](#)

1. Introduction

1. Summary

현대의 멀티미디어는 다양한 형태로 제공되고 있고
이러한 멀티미디어를 사용함에 있어
디지털 취약계층과 청각장애우의 제한된 미디어 접근성을 개선하고자 한다.

2. Introduce

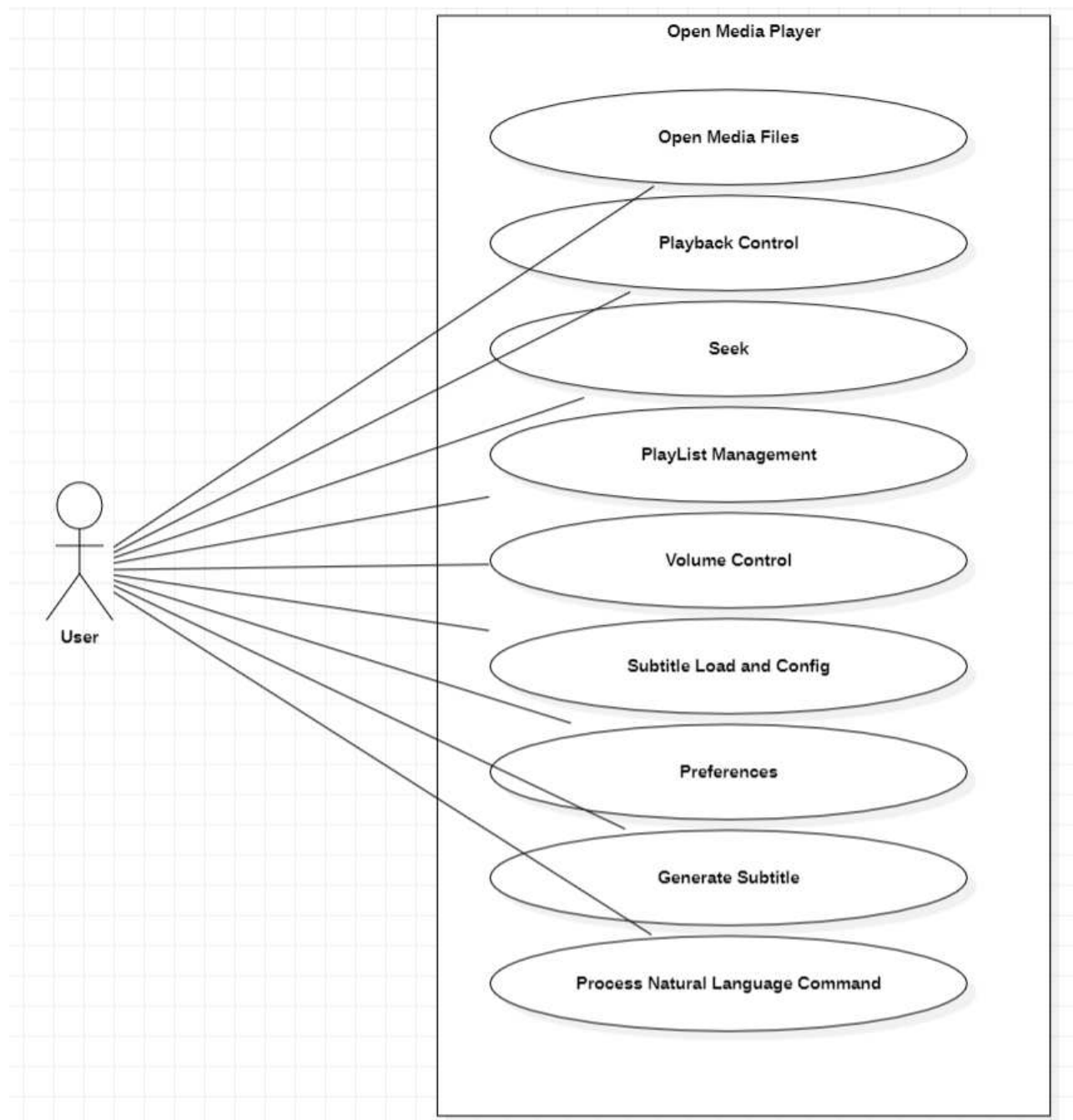
이번에 제작하게 된 Project, Open Media Player는
단순히 재생만 하는것이 아닌, STT를 이용한 영상의 자막생성과
LM을 사용한 Helper Chat-bot과 동작 자동화를 수행함으로써
청각장애우의 제한된 사용경험을 보완하고
디지털 취약계층의 접근성을 개선을 할 수 있을것으로 기대된다.

3. Goal

본 Analysis 문서에서 사용자가 어떤식으로 상호작용 하는지 분석하고
각 기능을 분석하여 명확하게 정의할 것이며.
이 작업을 수행함으로써, Open Media Player 프로젝트가 어떤식으로 구성되고
동작하게 되는지 상세하게 알 수 있을 것으로 기대된다.

2. Use case analysis

2-1. Use-Case Diagram



Open Media Player 시스템의 Use-Case Diagram을 나타낸 것이다.

Conceptualization에서 정의했던 Use-Case를 바탕으로 Diagram을 도출해 냈으며 Actor는 User 1개 도출되었다.

2-2. Use-Case Description

Use Case #1 : Open Media Files	
General Characteristics	
Summary	사용자가 파일 열기 메뉴, 버튼 또는 자연어 명령을 통해 재생할 미디어 파일을 선택하고 불러온다. 시스템은 해당 파일 형식과 필요한 코덱을 확인해 재생할 수 있도록 준비한다.
Scope	Open Media Player
Level	User
Author	BAE WON IL
Last Update	2025.05.07
Status	Analysis Phase (Draft)
Primary Actor	User
Preconditions	Open Media Player Application이 실행 중.
Trigger	사용자가 파일 열기 기능을 활성화 한다.
Success Post Condition	선택된 미디어 파일을 확인하고, 불러와 재생준비 완료 미디어 정보가 표시될 수 있음.
Failed Post Condition	오류메시지 표시 (Unsupported File Format, Couldn't Find File, Codec Error, etc) 파일 로드를 취소하고, 플레이어는 이전 상태 유지
Main Success Scenario	
Step	Action
1	사용자가 파일 열기 기능을 활성화 한다.
2	시스템이 파일 선택 대화상자를 표시한다.
3	사용자가 파일 시스템을 탐색해 미디어 파일을 선택한다.
4	사용자가 선택을 확인한다. (ex. 열기 버튼 누름)
5	시스템이 파일 형식과 코덱 유효성 검사
6	시스템이 파일 초기 부분을 디코드하고 재생 스트림 준비
7	시스템이 미디어 정보 표시하고, 재생준비 완료되었음을 표시 (ex. 재생 컨트롤 활성화, 첫 프레임 표시)
Extension Scenarios	
Step	Branching Action
5	5a. 지원하지 않는 파일 형식 5a.1. 시스템이 지원 하지 않는 파일 형식 오류 메시지 표시 5b. 필요한 코덱이 없음 5b.1. 시스템이 필요한 코덱이 없습니다 오류 메시지 표시
6	6a. 파일 손상 또는 읽기 불가 6a.1. 시스템이 파일 읽기 오류 메시지 표시
Related Information	
Performance	<= 10 sec (Typ <= 1s, 대용량 파일의 경우 10sec 이내 목표)
Frequency	
<Concurrency>	
Due Date	

Use Case #2 : Playback Control	
General Characteristics	
Summary	사용자가 GUI의 재생/일시정지/정지 버튼, 단축키, 또는 자연어 명령을 사용하여 현재 로드된 미디어의 재생 상태를 제어한다. 재생 명령 시 미디어를 출력하고, 일시정지 시 현재 재생 위치를 기록하며, 정지 시 관련 스트림 및 버퍼를 해제한다.
Scope	Open Media Player
Level	User
Author	BAE WON IL
Last Update	2025.05.07
Status	Analysis Phase (Draft)
Primary Actor	User
Preconditions	미디어 파일이 플레이어에 성공적으로 로드되어 있다.
Trigger	사용자가 '재생', '일시정지', 또는 '정지' 기능을 활성화한다.
Success Post Condition	(Play) - 미디어 재생이 현재 위치에서 시작되거나 재개된다. UI가 재생 상태를 반영한다. (Pause) - 미디어 재생이 현재 위치에서 일시 정지된다. UI가 일시 정지 상태를 반영한다. 현재 타임스탬프가 저장된다. (Stop) - 미디어 재생이 정지되고, 재생 위치는 처음으로 재설정된다. 관련 리소스가 해제될 수 있다. UI가 정지 상태를 반영한다.
Failed Post Condition	재생 중 오류(ex: 디코딩 오류) 발생 시 오류메시지 표시, 재생 정지
Main Success Scenario	
Step	Action
1	사용자가 '재생' 버튼을 클릭한다.
2	시스템이 현재 위치부터 오디오/비디오 프레임 디코딩 및 렌더링을 시작한다.
3	시스템이 UI를 업데이트하여 '재생 중' 상태와 재생 진행률을 표시한다.
4	사용자가 '일시정지' 버튼을 클릭한다.
5	시스템이 디코딩/렌더링을 중단하고 현재 타임스탬프를 저장한다.
6	시스템이 UI를 업데이트하여 '일시 정지됨' 상태를 표시한다.
Extension Scenarios	
Step	Branching Action
2	2a. 미디어 끝 도달 2a.1. 시스템이 자동으로 재생 정지. UI가 정지 또는 초기상태로 업데이트(다음 재생항목 재생 시) 2b. 재생 오류 2b.1. 시스템이 재생 정지, 오류메시지 표시
4	4a. 일시정지 대신 정지 4a.1. 시스템이 디코딩/렌더링 중단, 리소스 해제, 위치 재설정, UI 정지상태로 갱신
Related Information	
Performance	<= 0.5 sec
Frequency	
<Concurrency>	
Due Date	

Use Case #3 : Seek	
General Characteristics	
Summary	사용자가 재생 진행률 표시줄(Progress bar)을 직접 조작하거나, 특정 시간 이동 단축키를 사용하여 현재 재생 중인 미디어의 원하는 시간 위치(Timestamp)로 즉시 이동한다.
Scope	Open Media Player
Level	User
Author	BAE WON IL
Last Update	2025.05.07
Status	Analysis Phase (Draft)
Primary Actor	User
Preconditions	미디어 파일이 로드되어 재생 중이거나 일시 정지 상태이다.
Trigger	사용자가 재생 진행률 표시줄과 상호작용(클릭/드래그)하거나, 탐색 단축키를 사용
Success Post Condition	재생 위치가 사용자가 지정한 타임스탬프로 이동한다. 재생 중이었다면 새 위치에서 재생이 계속되고, 일시 정지 상태였다면 새 위치에서 일시 정지 상태를 유지한다. UI가 새 위치를 반영한다.
Failed Post Condition	탐색 작업 실패 시(예: 인덱스 없음, 잘못된 위치) 재생은 이전 위치에서 계속되거나 중지될 수 있다. 오류 메시지 표시
Main Success Scenario	
Step	Action
1	사용자가 재생 진행률 표시줄의 핸들을 클릭하거나 드래그하여 새 위치로 옮긴다.
2	시스템이 선택된 위치에 해당하는 목표 타임스탬프를 계산한다.
3	시스템이 내부 재생 포인터를 목표 타임스탬프로 이동시킨다.
4	시스템이 새 타임스탬프부터 디코딩 및 렌더링을 시작/재개한다.
5	시스템이 UI(진행률 표시줄, 표시 시간)를 업데이트하여 새 위치를 반영한다.
Extension Scenarios	
Step	Branching Action
3	3a. 탐색 목표가 미디어 경계 밖(ex. 음수 시간, 총 길이 초과) 3a.1. 시스템이 가장 가까운 유효 경계 지점으로 이동 3b. 미디어 탐색 불가 3b.1. 시스템이 탐색 요청을 취소하거나 오류 표시
Related Information	
Performance	<= 0.5 sec
Frequency	
<Concurrency>	
Due Date	

Use Case #4 : Playlist Management	
General Characteristics	
Summary	사용자가 인터페이스를 통해 여러 미디어 파일을 목록에 추가, 삭제하거나 순서를 변경하여 재생 목록을 생성, 저장 및 관리한다. 저장된 재생 목록을 불러와 순차, 반복, 또는 무작위 순서로 재생을 제어할 수 있다.
Scope	Open Media Player
Level	User
Author	BAE WON IL
Last Update	2025.05.07
Status	Analysis Phase (Draft)
Primary Actor	User
Preconditions	Open Media Player 애플리케이션이 실행 중이다.
Trigger	사용자가 재생 목록 인터페이스와 상호작용(파일 추가/삭제/순서 변경, 저장/로드 등)한다.
Success Post Condition	재생 목록이 사용자의 작업에 따라 수정(항목 추가/삭제/재정렬), 저장 또는 로드된다. 재생 목록 기반의 재생이 시작될 수 있다.
Failed Post Condition	작업 실패 시(ex. 재생 목록 저장 불가, 파일 추가 불가) 오류 메시지가 표시된다. 재생 목록은 이전 상태를 유지하거나 부분적으로 변경될 수 있다.
Main Success Scenario	
Step	Action
1	사용자가 재생 목록 창/패널을 연다.
2	사용자가 '파일 추가' 버튼을 통해 미디어 파일을 재생 목록에 추가한다.
3	시스템이 선택된 파일들을 재생 목록 뷰에 추가한다.
4	사용자가 재생 목록의 항목을 더블 클릭하거나 '재생' 버튼을 클릭한다.
5	시스템이 선택된/첫 번째 항목을 로드하고 재생을 시작한다 (Use Case #1, #2 트리거).
Extension Scenarios	
Step	Branching Action
2	2a. 사용자가 폴더 추가 2a.1. 시스템이 폴더를 확인해, 지원되는 미디어 파일 추가
3	3a. 일부 파일 유효하지 않거나 지원되지 않음 3a.1. 시스템이 유효한 파일만 추가하고, 유효하지 않은 파일은 무시
4	4a. 사용자가 재생목록 저장 4a.1. 시스템이 파일이름/위치를 요청하고 사용자가 정보제공, 시스템이 재생목록 파일 저장 4b. 사용자가 재생목록 불러옴 4b.1. 시스템이 재생 목록 파일을 물어봄, 사용자가 파일 선택, 시스템이 파일 참조 불러와 재생 목록 뷰에 표시
Related Information	
Performance	<= 1s 대규모 목록 로딩 시 UI응답성 유지
Frequency	
<Concurrency>	
Due Date	

Use Case #5 : Volume Control	
General Characteristics	
Summary	사용자가 GUI의 볼륨 슬라이더, 음소거 버튼, 단축키를 사용하여 시스템의 오디오 출력 레벨을 제어하거나 음소거 상태를 변경한다.
Scope	Open Media Player
Level	User
Author	BAE WON IL
Last Update	2025.05.07
Status	Analysis Phase (Draft)
Primary Actor	User
Preconditions	Open Media Player 애플리케이션이 실행 중이다.
Trigger	사용자가 볼륨 컨트롤(슬라이더, 음소거 버튼)과 상호작용하거나, 관련 단축키 사용한다.
Success Post Condition	오디오 출력 볼륨 레벨이 지정된 수준으로 조절되거나 음소거 설정. UI가 새 볼륨 상태를 반영한다.
Failed Post Condition	볼륨 조절 실패 시(ex. 오디오 장치 없음) 상태는 변경되지 않는다.
Main Success Scenario	
Step	Action
1	사용자가 볼륨 슬라이더를 원하는 레벨로 드래그한다.
2	시스템이 애플리케이션의 오디오 출력 레벨을 그에 맞게 조절한다.
3	시스템이 볼륨 슬라이더 위치를 업데이트하고 숫자 백분율을 표시할 수 있다.
Extension Scenarios	
Step	Branching Action
1	1a. 사용자가 음소거 클릭. 1a.1. 시스템이 볼륨을 0으로 설정, UI는 음소거 상태 표시 1b. 음소거 상태에서 음소거 버튼 클릭 1b.1. 시스템이 이전 볼륨 레벨 복원, UI는 음소거 해제 상태 표시
Related Information	
Performance	<= 0.2 sec
Frequency	
<Concurrency>	
Due Date	

Use Case #6 : Subtitle Load / Config	
General Characteristics	
Summary	사용자가 메뉴, 버튼을 통해 외부 자막 파일(SMI, SRT 등)을 로드하거나 미디어 내장 자막을 선택하여 화면에 표시한다. 자막의 표시 여부, 언어, 크기, 위치, 폰트, 색상, 싱크(시간 보정) 등을 조정한다.
Scope	Open Media Player
Level	User
Author	BAE WON IL
Last Update	2025.05.07
Status	Analysis Phase (Draft)
Primary Actor	User
Preconditions	비디오 미디어 파일이 로드되어 있다.
Trigger	사용자가 재생 목록 인터페이스와 상호작용(파일 추가/삭제/순서 변경, 저장/로드 등)한다.
Success Post Condition	선택된 외부 또는 내부 자막 트랙이 로드/선택되어 비디오와 동기화되어 표시된다. 자막 표시 설정(글꼴, 크기, 색상, 싱크 등)이 적용된다.
Failed Post Condition	자막 파일을 로드/파싱할 수 없거나 설정을 적용할 수 없다. 자막을 불러오지 않고, 경고 문구만 별도로 표시한다.
Main Success Scenario	
Step	Action
1	사용자가 '자막 파일 로드' 기능을 활성화한다.
2	시스템이 파일 선택 대화상자를 표시한다.
3	사용자가 자막 파일(예: .srt, .smi)을 선택한다.
4	사용자가 선택을 확인한다.
5	시스템이 자막 파일을 파싱한다.
6	시스템이 자막 타이밍을 비디오 재생 타임스탬프와 동기화한다.
7	시스템이 현재 타임스탬프 및 설정에 따라 적절한 자막 텍스트를 비디오 출력 영역에 표시.
Extension Scenarios	
Step	Branching Action
5	5a. 자막파일이 유효하지 않거나, 손상됨 5a.1. 시스템이 경고 문구를 표시하고, 로드하지 않음.
7	7a. 사용자가 자막 설정 열기 7a.1. 시스템이 설정옵션 대화상자를 표시한다. 7b. 사용자가 싱크 조절 7b.1. 시스템이 자막 표시 타임 스탬프에 시간 오프셋 적용 7c. 비디오에 내부 자막트랙 여럿 존재 7c.1. 사용자가 트랙을 선택하면 시스템이 선택된 트랙을 표시한다.
Related Information	
Performance	<= 0.1 sec
Frequency	
<Concurrency>	
Due Date	

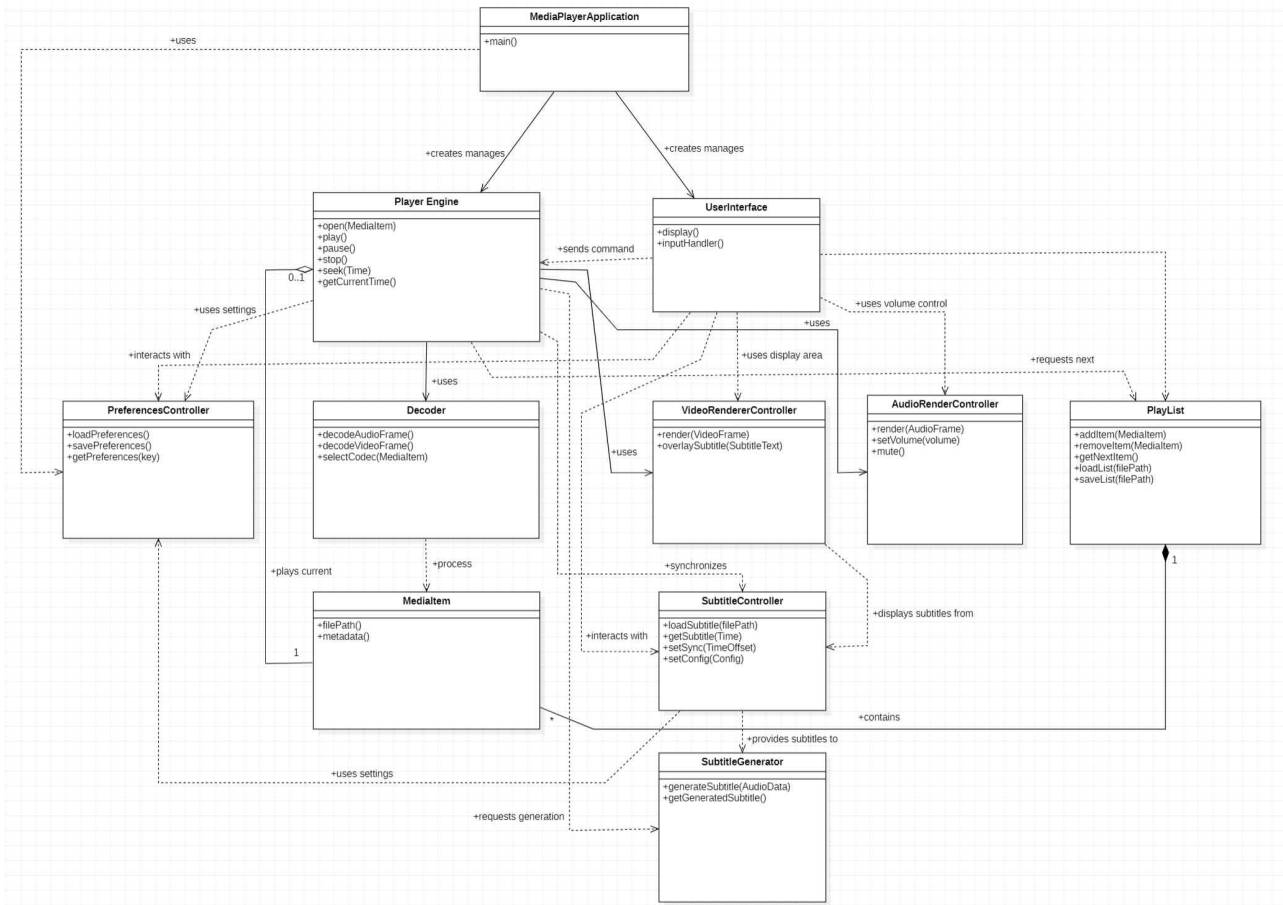
Use Case #7 : Preferences	
General Characteristics	
Summary	<p>사용자가 설정 메뉴를 열어 플레이어의 전반적인 동작 및 인터페이스 관련 설정을 확인하고 변경한다.</p> <p>설정 항목에는 기본 재생 옵션, 단축키 사용자 정의, 인터페이스 테마, 파일 연결, 코덱/필터 관리 등이 포함될 수 있다.</p> <p>변경된 설정은 저장되어 다음 실행 시에도 유지된다.</p>
Scope	Open Media Player
Level	User
Author	BAE WON IL
Last Update	2025.05.07
Status	Analysis Phase (Draft)
Primary Actor	User
Preconditions	Open Media Player 애플리케이션이 실행 중이다.
Trigger	사용자가 '환경 설정' 또는 '설정' 메뉴 옵션을 선택한다.
Success Post Condition	<p>환경 설정 창이 표시된다. 사용자가 설정을 수정하고 확인 ('확인', '적용' 등)하면 설정이 저장되고</p> <p>즉시 또는 플레이어 재시작 시 적용된다.</p>
Failed Post Condition	<p>설정을 저장할 수 없는 경우(ex. 권한 문제) 오류 메시지가 표시된다.</p> <p>설정은 변경되지 않는다.</p>
Main Success Scenario	
Step	Action
1	사용자가 '환경 설정' 기능을 활성화한다.
2	시스템이 다양한 설정 카테고리(ex. 재생, 인터페이스, 자막, 단축키)를 포함하는 환경 설정 창/대화상자를 표시한다.
3	사용자가 카테고리를 탐색하고 원하는 설정을 수정한다.
4	사용자가 변경 사항을 확인한다(ex. 적용 또는 확인 클릭).
5	시스템이 변경된 설정을 검증하고 저장한다.
6	시스템이 적용 가능한 설정을 적용한다.
7	시스템이 환경 설정 창을 닫는다 ('확인'을 클릭한 경우).
Extension Scenarios	
Step	Branching Action
4	<p>4a. 사용자가 취소 클릭</p> <p>4a.1. 시스템이 현재 변경된 모든 사항을 폐기하고, 창을 닫는다.</p>
5	<p>5a. 유효하지 않은 설정 값 입력</p> <p>5a.1. 시스템이 오류 메시지를 표시하고, 설정을 적용하지 않는다.</p>
Related Information	
Performance	<= 0.1 sec
Frequency	
<Concurrency>	
Due Date	

Use Case #8 : Generate Subtitle	
General Characteristics	
Summary	<p>사용자가 현재 재생 중인 비디오 미디어에서 자막 텍스트를 생성하도록 요청한다.</p> <p>생성된 자막은 현재 재생 시간과 동기화되어 화면에 표시된다.</p> <p>사용자는 STT 엔진의 종류나 인식 언어 등을 선택할 수 있다.</p>
Scope	Open Media Player
Level	User
Author	BAE WON IL
Last Update	2025.05.07
Status	Analysis Phase (Draft)
Primary Actor	User
Preconditions	비디오 파일이 로드되어 있다.
Trigger	사용자가 '자막 생성' 버튼/메뉴 옵션을 클릭한다.
Success Post Condition	<p>시스템이 오디오를 추출하고, STT를 수행하며, 타임스탬프가 있는 자막 데이터를 생성하고, 비디오와 동기화하여 표시한다.</p> <p>생성된 자막은 저장 가능할 수 있다.</p>
Failed Post Condition	<p>오디오 추출 실패, STT 처리 실패.</p> <p>오류 메시지가 표시된다.</p> <p>자막이 생성되지 않는다.</p>
Main Success Scenario	
Step	Action
1	사용자가 '자막 생성' 기능을 활성화한다.
2	시스템이 사용자에게 작업을 확인한다
3	시스템이 비디오 파일에서 오디오 트랙 추출을 시작한다.
4	시스템이 추출된 오디오를 설정된 STT 엔진에 공급한다.
5	STT 엔진이 오디오를 처리하고 타임스탬프가 있는 텍스트 세그먼트를 반환한다.
6	시스템이 생성된 자막을 비디오 재생과 동기화하여 표시한다.
Extension Scenarios	
Step	Branching Action
3	<p>3a. 오디오 추출 오류</p> <p>3a.1. 시스템이 오류 메시지를 표시한다.</p>
7	<p>7a. 사용자가 생성된 자막 선택</p> <p>7a.1. 시스템이 파일 이름/위치를 요청, 사용자가 정보 제공, 시스템이 제공된 정보로 파일 저장</p>
Related Information	
Performance	<p><= 10s</p> <p>STT처리는 시간이 오래 걸릴 수 있으므로, 일정 시간(15초) 이상 소요 시</p> <p>사용자에게 피드백을 제공.</p>
Frequency	
<Concurrency>	
Due Date	

Use Case #9 : Process Natural Language Command	
General Characteristics	
Summary	<p>사용자가 텍스트 입력 인터페이스를 통해 플레이어 제어나 정보 요청을 자연어 형태로 입력한다.</p> <p>시스템은 이 명령의 의도를 파악하여, 플레이어의 특정 기능 (ex. 재생 제어, 볼륨 조절, 특정 설정 메뉴 열기, 자막 생성 시작 등)을 자동으로 실행한다.</p> <p>(ex. 설정열어줘 -> 환경설정 열기, 자막 만들어줘 -> 기본설정으로 자막 생성)</p>
Scope	Open Media Player
Level	User
Author	BAE WON IL
Last Update	2025.05.07
Status	Analysis Phase (Draft)
Primary Actor	User
Preconditions	Open Media Player 애플리케이션이 실행 중이다.
Trigger	<p>사용자가 Help-Live Support메뉴를 선택해</p> <p>지정된 텍스트 입력 영역에 명령을 입력한다.</p>
Success Post Condition	<p>SLM이 사용자 명령의 의도와 매개변수를 정확히 해석한다.</p> <p>시스템이 해당하는 작업(예: 재생 시작, 볼륨 변경, 설정 창 열기)을 실행한다.</p> <p>피드백이 제공될 수 있다.</p>
Failed Post Condition	<p>SLM이 명령을 이해하지 못하거나, 요청된 작업을 수행할 수 없다(부적절한 요청).</p> <p>실패를 나타내는 피드백 (ex. 이해하지 못했습니다, 해당 작업을 수행할 수 없습니다)이 제공된다.</p> <p>시스템 상태는 변경되지 않는다.</p>
Main Success Scenario	
Step	Action
1	사용자가 명령 입력을 활성화하고 자연어 요청을 입력한다.
2	Local SLM이 텍스트를 분석하여 의도와 매개변수를 식별한다
3	의도와 매개변수를 시스템 매크로 함수에 인자로 넘겨서 실행한다
4	시스템이 해당 기능을 실행하여 사용자에게 제공한다.
5	시스템이 확인 피드백을 제공할 수 있다.
Extension Scenarios	
Step	Branching Action
2	<p>2a. 명령 이해 불가</p> <p>2a.1. 시스템이 response로 명령을 이해하지 못했다는 피드백 제공</p> <p>ex. “죄송합니다, 이해하지 못했습니다, 무엇이 필요한지 다시 설명해 주시겠습니까?”</p>
	<p>2b. 부적절한 프롬프트</p> <p>2b.1 시스템이 response로 해당 작업을 수행할 수 없다는 피드백 제공</p> <p>ex. “죄송합니다, 해당 요청은 수행할 수 없습니다”</p>
4	<p>4a. 현재 상태에서 요청 작업 불가</p> <p>7a.1. 시스템이 해당 요청에 대해 response로 피드백 제공</p> <p>ex. “현재 상태에서는, 해당 작업을 수행할 수 없습니다.”</p>
Related Information	

Performance	<= 5 sec 요청 제공 후 작업 수행까지 반응시간 최소화
Frequency	
<Concurrency>	
Due Date	

3. Domain analysis



위 그림은 본 프로젝트의 Domain Diagram을 표시한 것이다.

1) MediaPlayerApplication

Application의 시작점이며, 전체 구성요소를 관리하는 최상위 클래스
UserInterface와 PlayerEngine을 생성 및 관리하며
PreferencesController를 사용해 초기 설정을 로드한다.

관리 – UserInterface, PlayerEngine

사용 – PreferencesController

2) UserInterface

사용자와의 GUI 상호작용 담당, 사용자 입력을 받아
PlayerEngine, AudioRenderController, Playlist, SubtitleController,
PreferencesController 등과 상호작용해 해당 기능 실행을 요청한다.
VideoRenderController가 비디오 표시 영역을 제공하고, 재생상태 및 정보를
표시한다.

상호작용 – PlayerEngine, Playlist, SubtitleController, PreferencesController

사용 – VideoRenderController, AudioRenderController

3) PlayerEngine

미디어 재생의 핵심 로직을 제어.

현재 재생할 Medialtem을 관리하며(Aggregation, 0..1 관계),

Decoder를 통해 미디어 데이터를 디코딩하고,

AudioRenderController와 VideoRenderController를 사용하여 오디오/비디오
출력을 관리합니다.

Playlist로부터 다음 재생 항목을 가져오고,

SubtitleController와 재생 시간을 동기화하며,

SubtitleGenerator에 자막 생성을 요청하고,

PreferencesController로부터 재생 관련 설정을 읽어올 수 있다.

집합 – Medialtem

사용 – Decoder, AudioRenderController, VideoRenderController

상호작용 – Playlist, SubtitleController, SubtitleGenerator, PreferencesController

4) Playlist

여러 Medialtem의 목록을 관리.
Medialtem 객체들을 소유(Composition, 1 대 * 관계)하며,
목록에 아이템을 추가/삭제하고,
재생 순서를 관리하며,
재생 목록 파일을 저장하고 로드하는 기능을 제공.
복합 – Medialtem

5) Medialtem

개별 미디어 파일(오디오 또는 비디오)을 나타내는 클래스.
파일 경로, 재생 시간, 제목 등의 메타데이터 정보를 포함.
Playlist에 의해 관리되고, Decoder에 의해 처리되며,
PlayerEngine에 의해 현재 재생 대상으로 참조.
포함 – Playlist
처리 – Decoder
참조 – PlayerEngine

6) Decoder

Medialtem으로부터 오디오 및 비디오 데이터를 읽어들이
시스템이 처리 가능한 원시 데이터(raw frame/sample)로 변환(디코딩).
적절한 코덱을 선택하고 관리.
처리 – Medialtem
사용 – PlayerEngine

7) AudioRenderController

PlayerEngine으로부터 디코딩된 오디오 데이터를 받아
오디오 출력 장치(스피커 등)를 통해 재생.
UserInterface 또는 PlayerEngine의 요청에 따라 볼륨 조절 및 음소거 수행.
사용 – PlayerEngine, UserInterface

8) VideoRenderController

PlayerEngine으로부터 디코딩된 비디오 프레임 데이터를 받아 UserInterface가
제공하는 화면 영역에 표시.
SubtitleController로부터 받은 자막 텍스트를 비디오 위에 오버레이하여 출력하는
기능을 수행.
사용 – PlayerEngine, UserInterface
상호작용 – SubtitleController

9) SubtitleController

자막 파일(SMI, SRT 등)을 로드, 파싱하고 관리.

PlayerEngine으로부터 현재 재생 시간을 받아 동기화된 자막 텍스트를 VideoRenderController에 제공.

사용자의 자막 관련 설정(싱크, 폰트, 크기 등)을 관리,

이 설정은 PreferencesController를 통해 저장/로드될 수 있음.

SubtitleGenerator로부터 생성된 자막 데이터를 받음.

상호작용 – PlayerEngine, VideoRenderController, SubtitleGenerator, UserInterface

사용 – PreferencesController

10) SubtitleGenerator

PlayerEngine 또는 Decoder로부터 오디오 데이터를 받아 처리하고,

생성된 타임스탬프 포함 자막 데이터를 SubtitleController에 전달.

상호작용 – PlayerEngine, SubtitleController

11) PreferencesController

애플리케이션의 설정값을 로드하고 저장하는 역할을 담당.

다양한 컴포넌트(PlayerEngine, SubtitleController, UserInterface 등)가 필요로 하는 설정 정보를 제공합니다.

사용 – MediaPlayerApplication, PlayerEngine, SubtitleController

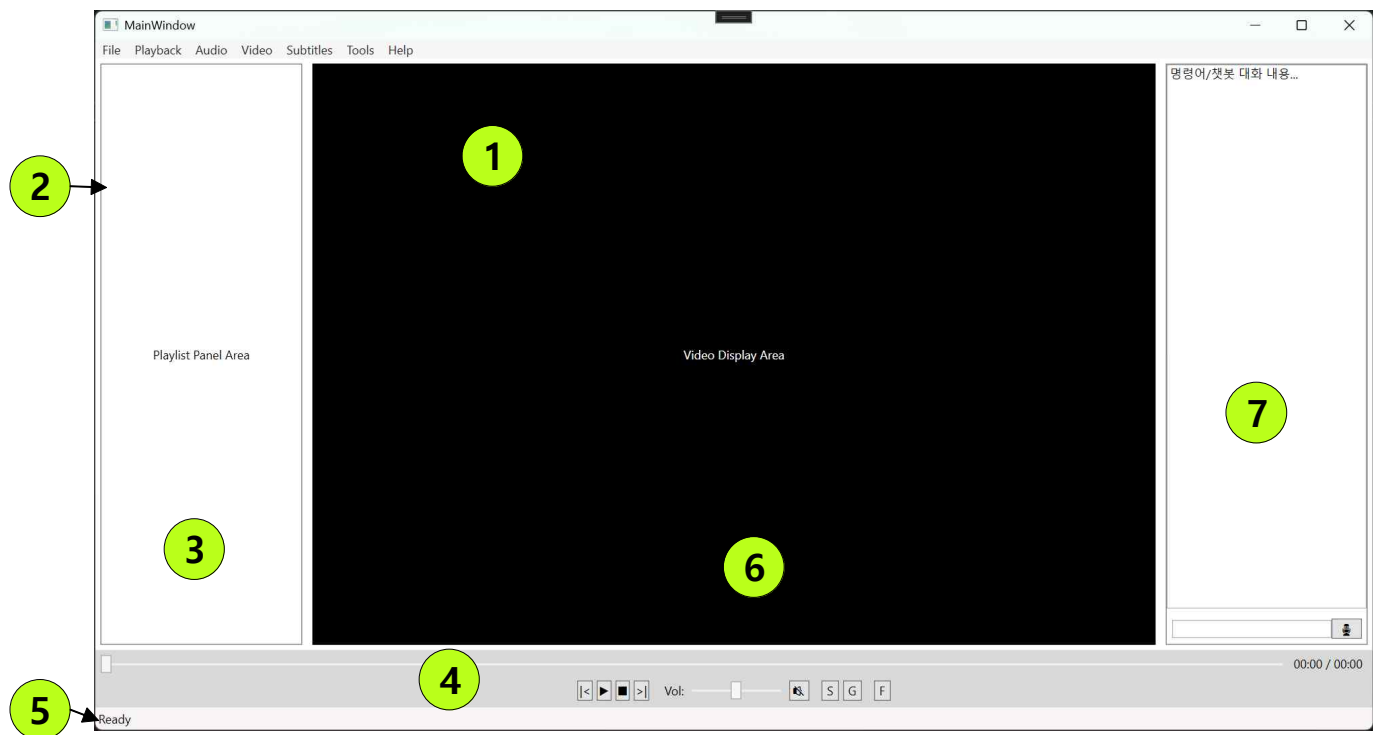
상호작용 – UserInterface

4. User Interface prototype

본 파트에서 위에 분석한 내용을 바탕으로 기본적인 prototype을 만들어 기본적인 UI구성을 확인한다.

현재 구성은 Prototype으로써, 개발중 수정되어 변경될 수 있음.

1.기본 화면(전체 호출)



1: Window Name Space : 현재 애플리케이션의 이름과 조작을 표시 (Min, Max, Close)

2: Menu Bar : 세부 조작 메뉴 표시

File : 파일관련 메뉴 (열기, 불러오기(자막))

Playback : 재생관련 메뉴 (재생, 일시정지, 정지, 배속)

Audio : 오디오 제어 관련 메뉴 (볼륨조정, 재생트랙지정)

Video : 비디오 제어 관련 메뉴 (화면캡처, 비디오 컬러 조정)

Subtitles : 자막 제어 관련 메뉴 (자막싱크, 자막생성, 자막설정)

Tools : 기타 설정 메뉴 (환경설정)

Help : 도움말 메뉴 (정보, Live Support)

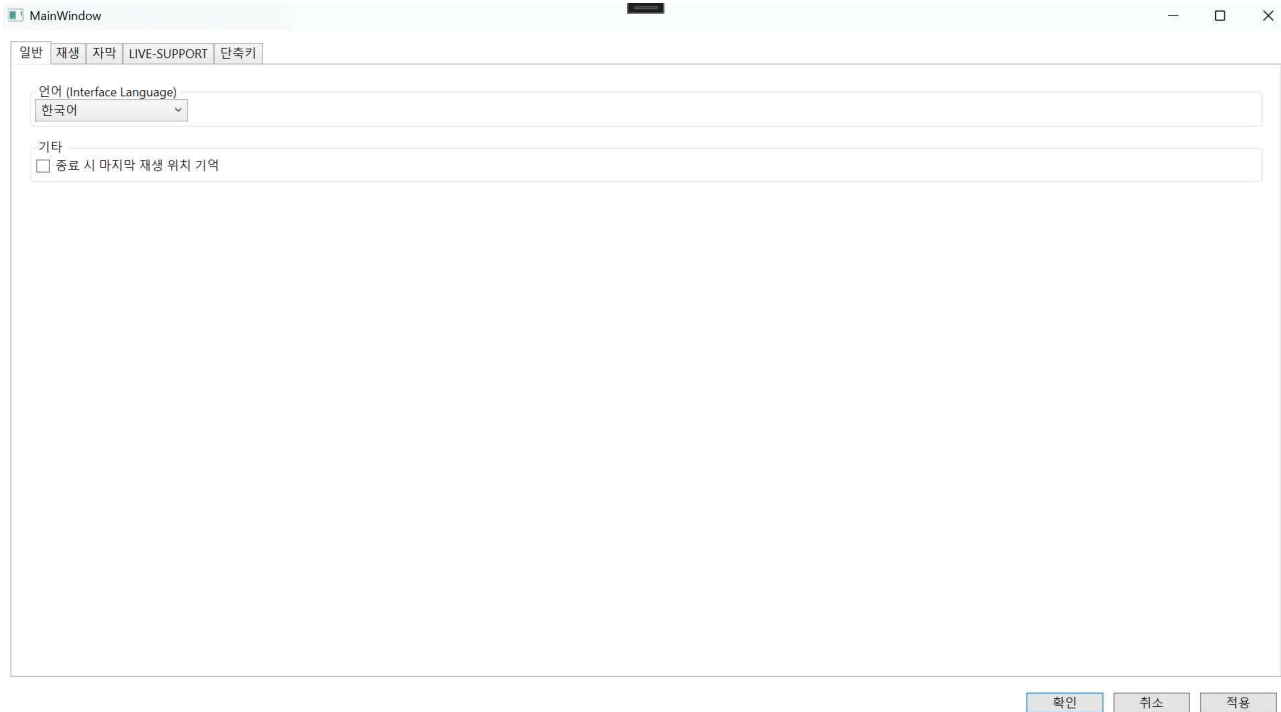
3: Playlist : 현재 플레이 리스트 표시

4: 재생바 : 현재 재생 정보 표시 및, 재생제어

(이전/재생-일시정지/정지/다음/볼륨제어/음소거/자막ON-OFF/자동자막생성/전체화면)

- 5: 현재 상태 표시 (READY/PLAYING/PAUSE/STOP/PROCESSING...)
- 6: 현재 재생중인 비디오 프레임 표시
- 7: LIVE-SUPPORT 메뉴 호출 시 표시

2. 설정 화면



MainWindow

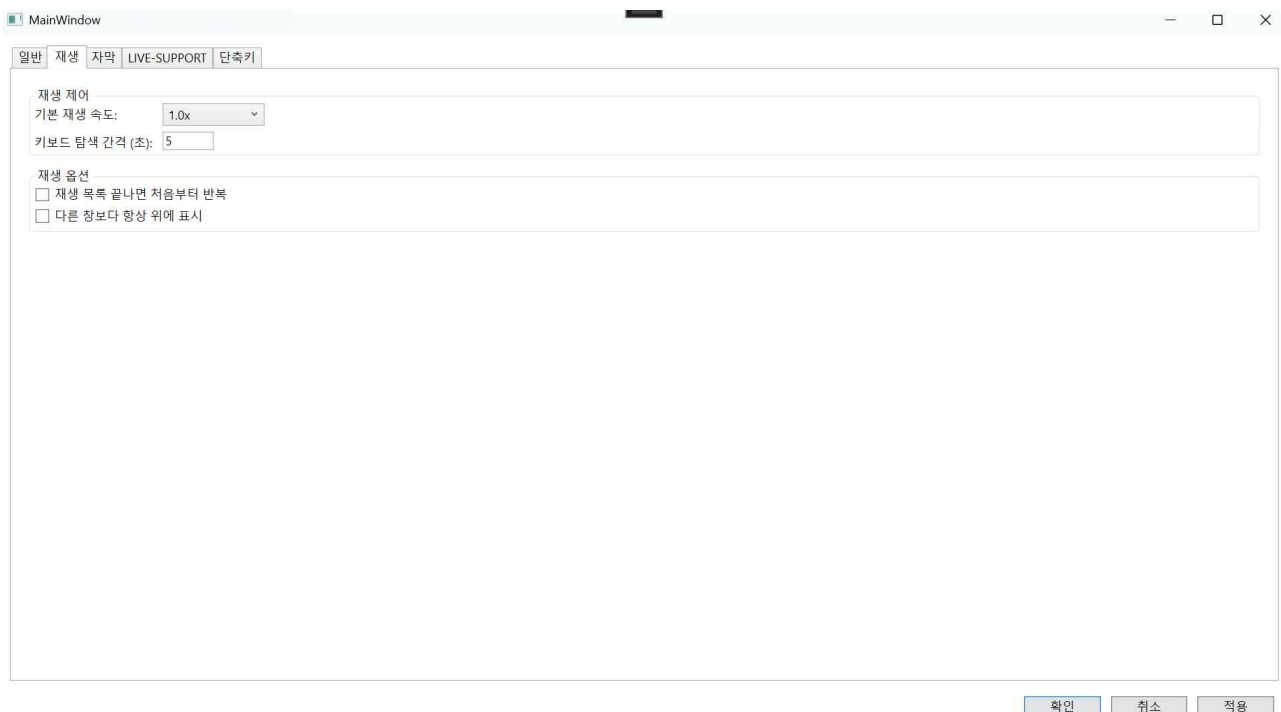
일반 재생 자막 LIVE-SUPPORT 단축키

언어 (Interface Language)
한국어

기타
☐ 종료 시 마지막 재생 위치 기억

확인 취소 적용

기본 설정, 언어설정과, 기본적인 설정을 저장한다



MainWindow

일반 재생 자막 LIVE-SUPPORT 단축키

재생 제어
기본 재생 속도: 1.0x
키보드 탐색 간격 (초): 5

재생 옵션
☐ 재생 목록 끝나면 처음부터 반복
☐ 다른 정보도 항상 위에 표시

확인 취소 적용

재생 설정, 미디어 재생시 기본 배속과 탐색 간격을 설정한다.
그 외 재생옵션도 필요에 따라 설정한다

MainWindow

일반
재생
자막
LIVE-SUPPORT
단축키

표시 설정 (Appearance)

기본 글꼴:
글꼴 선택...

글꼴 크기 배율 (%):

글꼴 색상:
색상 선택...

배경/윤곽선:
없음

표시 위치 (%):

로드 및 동기화

기본 인코딩:
자동 감지 (CP949/UTF-8)

선호 언어 (원표 구분):
ko,en

기본 싱크 (초):
0.0

자동 생성 (STT) 설정

STT 엔진/모델:
Whisper (Local) - Base

음성 인식 언어:
자동 감지

☐ 비디오 파일 열 때 자동으로 자막 생성 시도

확인
취소
적용

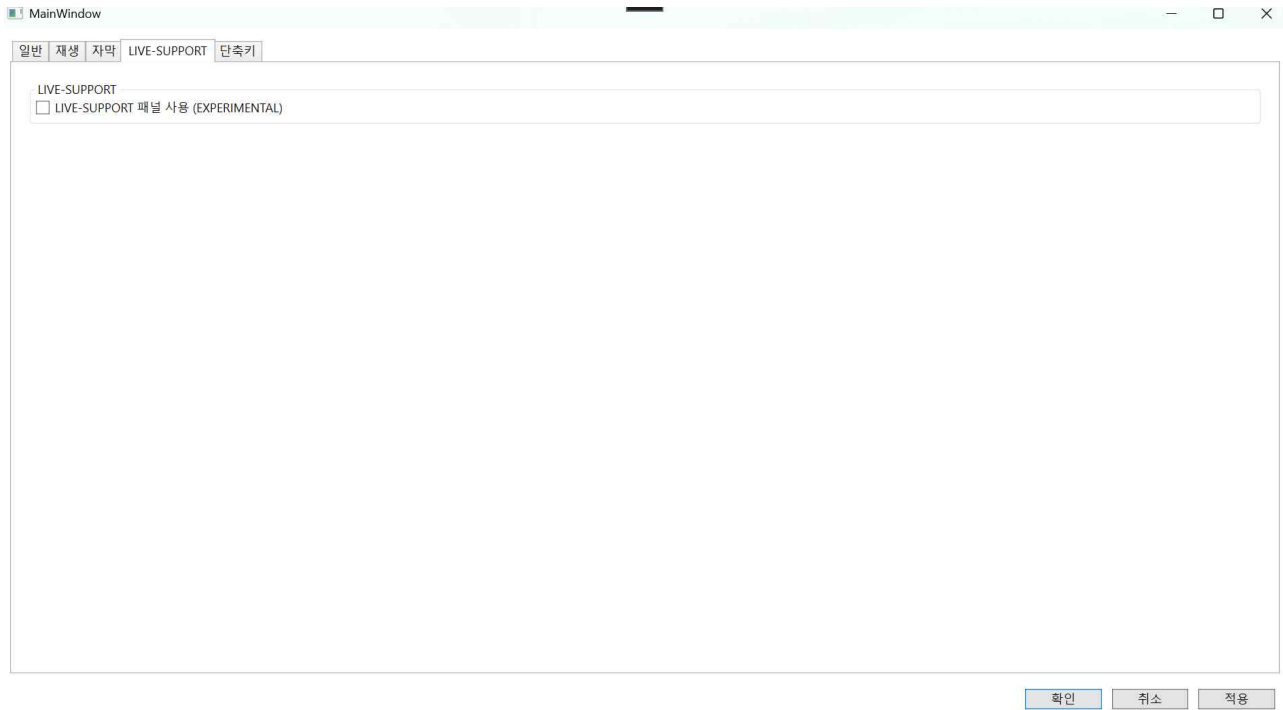
자막표시 및 설정 관련

자막 표시 설정을 지정할 수 있다.

로드 및 동기화 메뉴에서 기본 인코딩 및 선호 인코딩 설정이 가능하다.

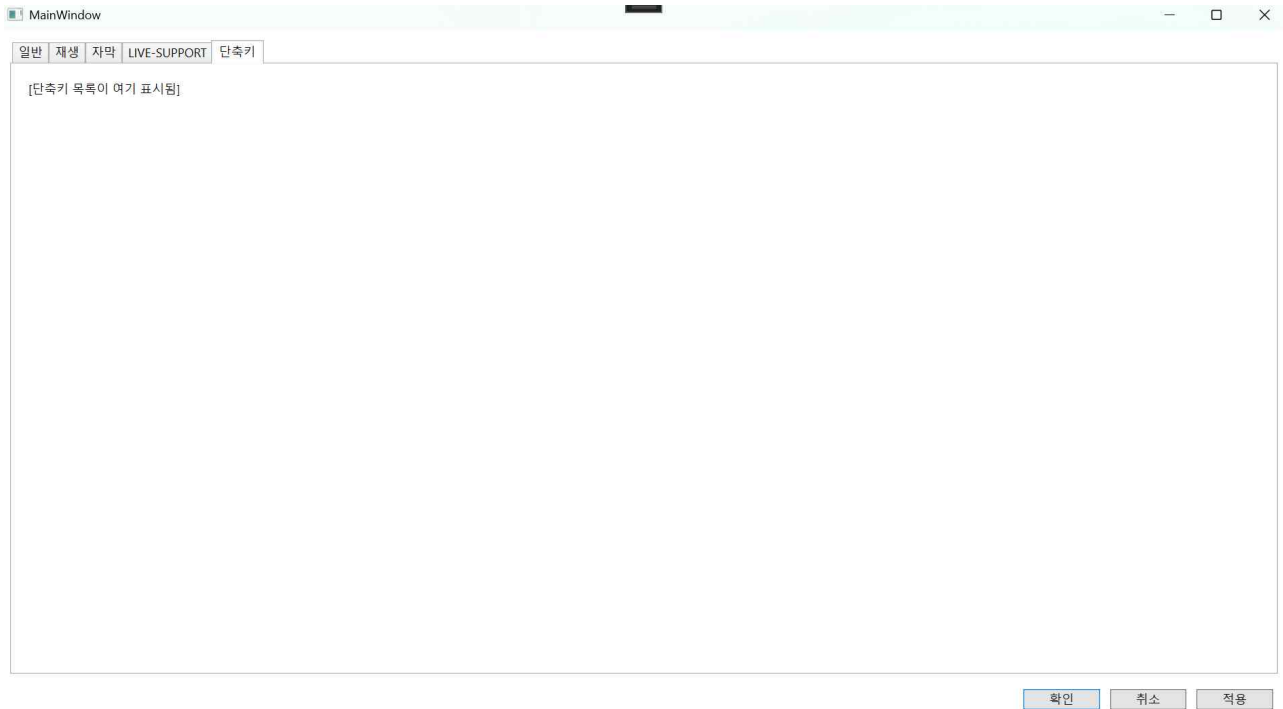
자동 생성 설정에서 stt엔진 옵션과 기본인식언어

비디오 open시 자동 생성 시도를 지정할 수 있다.



LIVE-SUPPORT

SLM을 통한 사용자 실시간 지원 설정이다, 현재는 live tab ON/OFF만 설정할 수 있다.



단축키를 표시하는 탭으로, 현재는 단순히 표시만 해주는 것으로 진행중이다.

5. Glossary

API	응용 프로그램이 운영체제 기능이나 다른 서비스의 기능을 사용할 수 있도록 미리 정의된 명령어, 함수, 프로토콜 등의 인터페이스 규격.
Chatbot (챗봇)	사용자와의 대화(텍스트 또는 음성)를 통해 정보를 제공하거나 특정 작업을 자동화하는 프로그램.
Playlist	여러 미디어 파일을 원하는 순서대로 모아놓은 목록
Codec	멀티미디어 데이터(오디오, 비디오)를 압축(Encoding)하거나 압축 해제(Decoding)하는 소프트웨어 또는 하드웨어 알고리즘입니다. (예: H.264, AAC, VP9).
Decoder	코덱을 사용하여 압축된 오디오 또는 비디오 데이터를 플레이어가 처리할 수 있는 원시(Raw) 데이터 형태로 변환하는 구성 요소입니다.
FFMPEG	오디오 및 비디오 파일의 기록, 변환, 스트리밍을 위한 매우 광범위한 라이브러리와 도구를 제공하는 오픈 소스 프로젝트.
LM / SLM	인간의 언어(자연어)를 통계적으로 모델링하여 이해하거나 생성할 수 있는 인공지능 모델. SLM은 비교적 적은 파라미터 수를 가진 경량화된 언어 모델을 의미.
Renderer	디코딩된 오디오 샘플이나 비디오 프레임 데이터를 실제 출력 장치(스피커, 화면)를 통해 사용자가 인지할 수 있도록 출력하는 구성 요소.
STT	사람의 음성 언어를 컴퓨터가 읽을 수 있는 텍스트 데이터로 변환하는 기술입니다.
Whisper	OpenAI에서 개발하여 공개한 오픈 소스 음성 인식(STT) 모델입니다. 높은 정확도와 다국어 지원 능력을 특징으로 합니다.

6. References

Open-AI Whisper(paper) : <https://arxiv.org/abs/2212.04356>

Open-AI Whisper(git) : <https://github.com/openai/whisper>

FFMPEG : <https://www.ffmpeg.org/documentation.html>

MS DirectX API: <https://learn.microsoft.com/en-us/windows/win32/directx>

MS Win32 API :

<https://learn.microsoft.com/en-us/windows/win32/apiindex/api-index-portal>

MS .NET API : <https://learn.microsoft.com/en-us/dotnet/api/>

WPF .NET:

<https://learn.microsoft.com/en-us/dotnet/desktop/wpf/overview/?view=netdesktop-9.0>

참고문헌

LLM-Powered GUI Agents in Phone Automation: Surveying Progress and Prospects

<https://arxiv.org/pdf/2504.19838>

B-MOCA:BENCHMARKING MOBILE DEVICE CONTROL AGENTS

ACROSS DIVERSE CONFIGURATIONS

<https://arxiv.org/pdf/2404.16660>

LLM-Powered GUI Agents in Phone Automation: Surveying Progress and Prospects

B-MOCA:BENCHMARKING MOBILE DEVICE CONTROL AGENTS ACROSS

DIVERSE CONFIGURATIONS

위 문헌들을 참고하여 추가된 부분이다.

LLM-Powered GUI Agents in Phone Automation에서

Language Model을 사용한 벤치마크 및 사용자 동작 자동화를 진행할 수 있음에 영감을 받았으며

국내 KAIST에서 진행했던 B-MOCA에서 Mobile-Benchmark를 진행함에 있어

Language Model을 사용해 기존의 정적인 매크로 보다 일반 사용자에게 있어

기반지식 없이도 자연어 요청을 통해 매크로 수행(자동화 동작)이 가능하다는 점에서 접근성을 개선 시킬 수 있을것으로 기대된다.

본 프로젝트에서는 별도 메뉴에서 사용자가 자연어로 요청하면

LM이 해당 요청을 확인하여 매크로 수행이 가능하도록 기능을 추출하고

추출된 파라미터를 매크로 함수에 인자로 넘겨주어 기능을 수행하도록 진행중이다.

Local에서 사용할 SLM은 DeepSeek-V3, Gemma2, Llama3.2를 선정하였고
아주 높은 성능이 필요한게 아닌 관계로 1~7b 사이의 작은 모델을 사용한다.
Fine-Tuning과 system Instruction(Message)사이에서 테스트 후
system instruction으로 충분한 성능이 나오는 경우
별도로 Fine-Tuning은 진행하지 않아도 될 것으로 보임.

마찬가지로 Whisper STT는 기본으로 Tiny를 사용 후 필요에 따라 여러 규모의 모델을
선택할 수 있도록 할 예정