



# Cloud Platform Project Idea

Team: 민초단

김지현 20160131  
조은정 20160192  
이연재 20170225

A sepia-toned photograph of a person's hands pouring coffee from a kettle into a white cup. Inside the cup, a miniature mountain landscape with a small hut is visible. The background is dark, and the overall mood is cozy and artistic.

# Contents

**01 주제 및 대상**

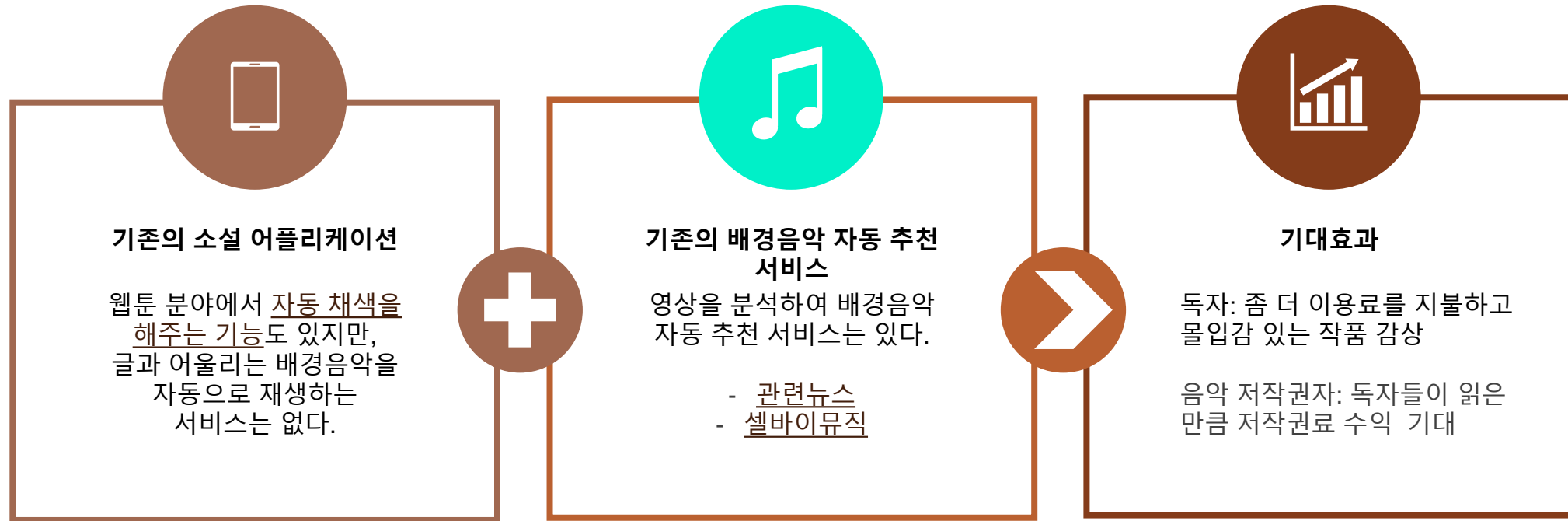
**02 데이터 수집 방법 및 상세**

**03 구현방식**

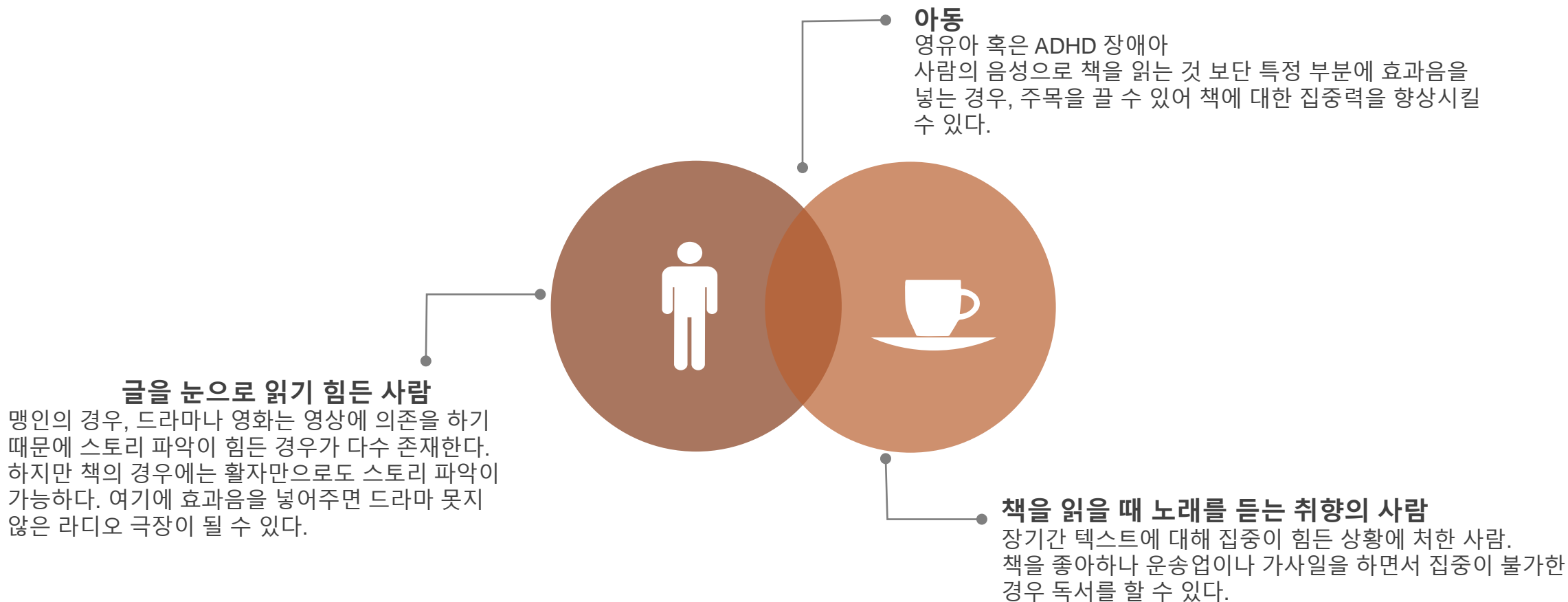
**04 Timeline과 Reference**

# 주제

## 글(소설)과 어울리는 BGM (배경음악) 매칭 재생



# Target



# 데이터 수집 방법



## 글 수집

- 구텐베르크 프로젝트
  - 1) 국내: <http://m.nextpub.net/>
  - 2) 해외: <https://www.gutenberg.org/>



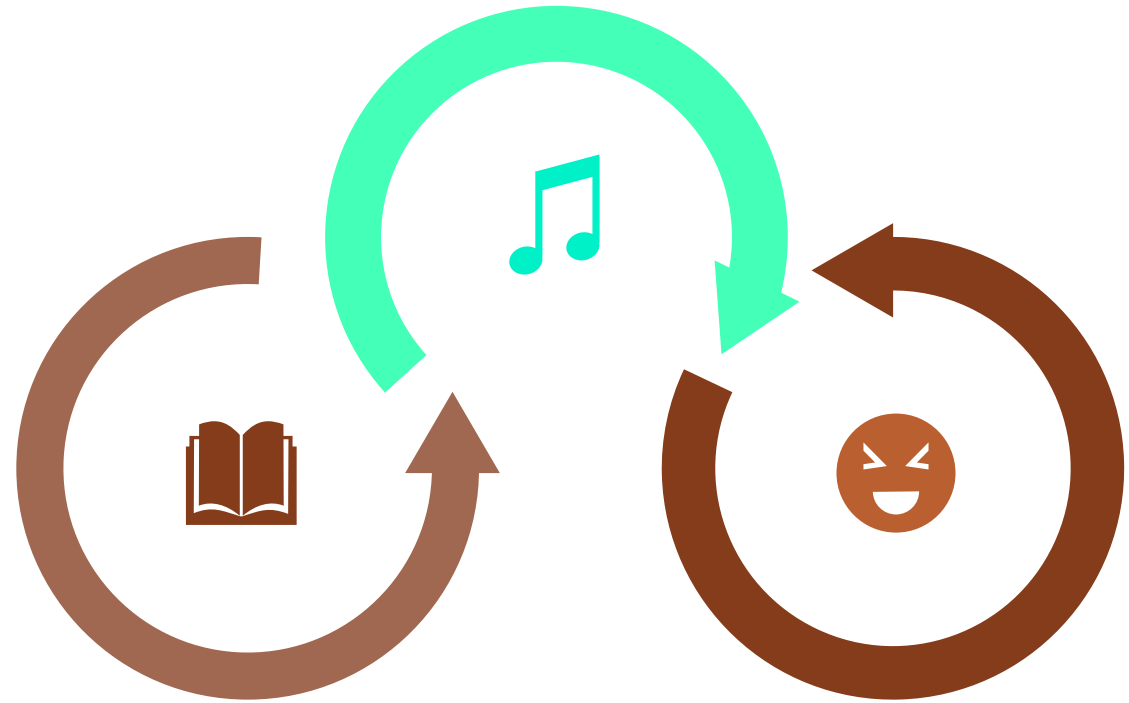
## 음악

브금저장소  
셀바이뮤직  
유튜브 플레이리스트 영상  
카카오 아레나 멜론 데이터



## 감정 사전

기존의 감성 사전이 구축되어 있는지 확인



# 데이터 수집 상세



Text Data

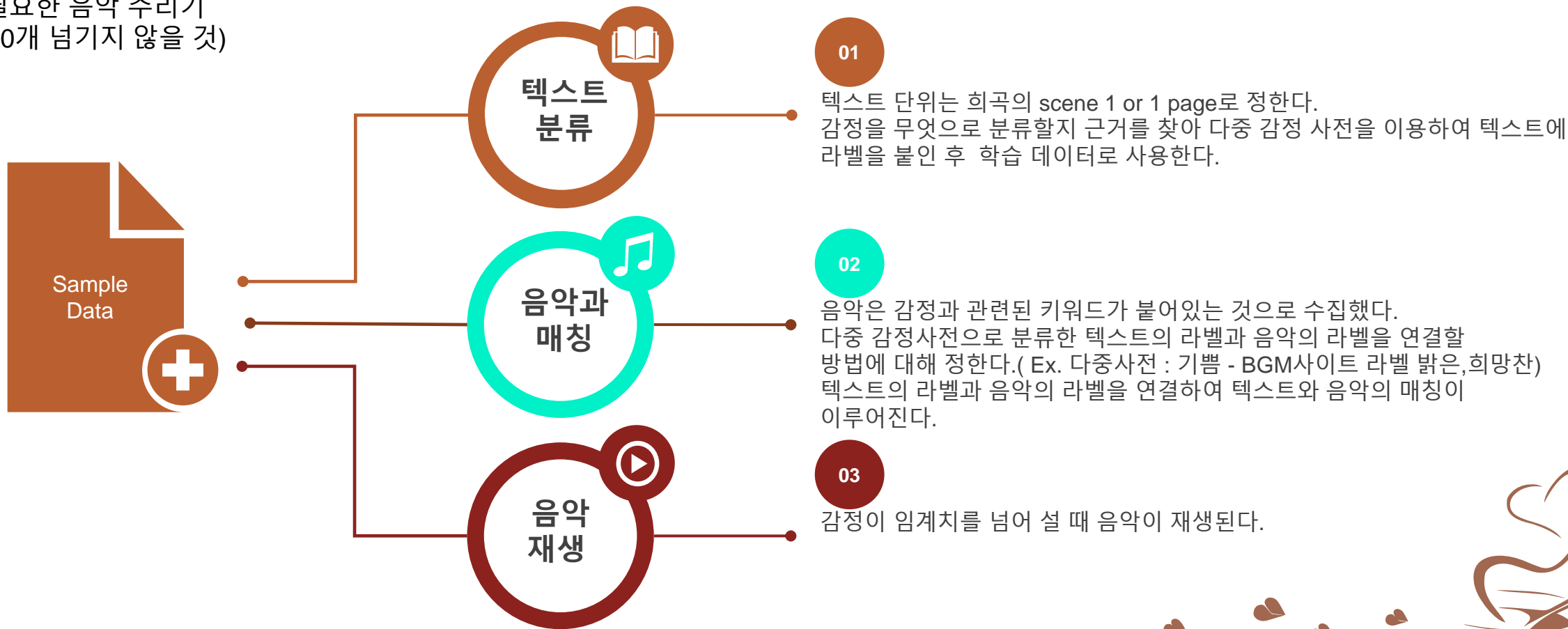


Music Data

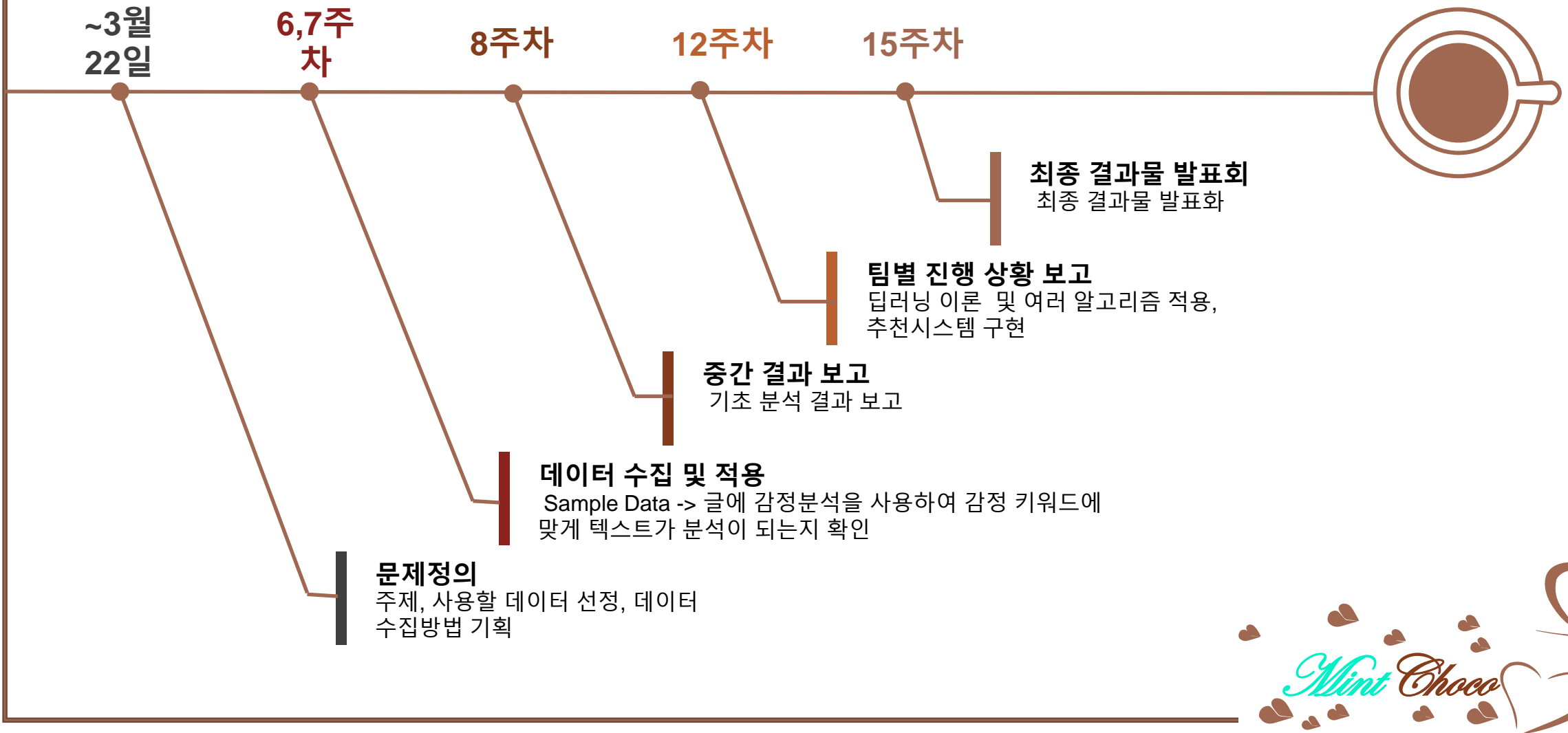
Sample(데이터 수집 시)	저작권여부	X	X
	장르	셰익스피어 작품, 동화	효과음, 클래식
	예시	로미오와 줄리엣, 한여름밤의 꿈, 백설공주, 이솝우화	
Service(상용화 시)	저작권여부	O/X	O/X
	장르	웹소설, 고전, 동화	클래식, 가요, 효과음

# 구현 방식

- 음악 데이터 수 정하기
    - 음악데이터는 오랜 시간 소요
    - 프로젝트 규모 정리하기
    - 확인용으로 돌리고 그 다음단계에 더 추가하기
    - 셰익스피어를 1차 목표로 정하기
    - 필요한 음악 추리기
- (1000개 넘기지 않을 것)



# Project Timeline





# Reference Paper



01

개인의 감성분석과 머신러닝을 적용한 노래 추천 서비스 논문

고객의 감정 분석, 상황 분석과 머신 러닝을 적용하여 현재 매장의 분위기와 어울리는 음악이 선곡되고 재생될 수 있도록 하는 것을 목표로 하고 있는 논문

02

텍스트 감정 분석과 음성 합성을 이용한 오디오북 서비스

텍스트에서 감정 정보를 분석하고 이를 기반으로 감정에 따른 음성 합성을 하여 사용자 친화적 오디오북을 제작하기 위한 서비스를 소개

03

주의력 결핍 및 과잉행동 문제를 가진 아동의 주의집중력 증진을 위한 음악인지프로그램의 효과 연구

소리의 공간적 및 시간적 속성 지각과 인지, 음악적 정보처리와 같은 인지기술 강화를 위해 구조화된 음악활동이 주의집중력 증진에 영향을 미친다.

04

SentencePiece 토큰라이저와 부분단어 임베딩 기법을 이용한 비형식적 텍스트의 감정 분석

워드 임베딩 기법의 단점으로 제시되는 OOV(Out-Of-Vocabulary) 문제를 해결하기 위해 부분단어(subword) 단위의 토큰화 기법인 SentencePiece를 이용한 감정 분석을 제안



Thank You

## 생각나는거 추가하기

텍스트 라벨링이 일찍 끝날 경우:

- 다른 장르의 텍스트 데이터 돌려보기 (ex. 앱소설)
- 음악 데이터를 분류할 때 키워드 말고 파동으로 분류하기(난이도 상)

시연결과물 예시:

- 어플 만들어서 결과물 시연
- 낭독하면서(좋아하는 목소리로) 음악도 나오도록

## 감성사전

The screenshot shows the BK21 Plus website. The top navigation bar includes links for '사업팀 소개' (Team Introduction), '자료실' (Data Room), '사업팀 구성' (Team Structure), '정보광장' (Information Plaza), '사업팀 현황' (Team Status), and '사업신청서' (Application Form). The main content area is titled '공유자료' (Shared Data) and displays a table of shared files. A file named '감성단어사전0603.xlsx (32.3K)' is highlighted with a red box. Below the table, there is a notice about the file's origin and a disclaimer.

제목	작성자	등록일	조회수
[빅데이터분석팀] 감성단어 목록	관리자	15-06-03 21:34	11,224
감성단어사전0603.xlsx (32.3K)	관리자	15-06-03 21:34	

감성단어사전0603.xlsx (32.3K) 12/100 DATE : 2015-06-03 21:34:45

김성민 수강자의 도움으로 감성단어 목록 파일을 만들었습니다.  
인용 자료는 다음과 같습니다.  
손신주, 박미숙, 박지훈, & 손진훈, (2012). 한국어 감성표현단어의 추출과 범주화, *감성과학*, 1(1), 105-120.  
프로젝트 수행에 도움이시기 바랍니다.

03/30 실시간 수업

텍스트 데이터 탐색

-