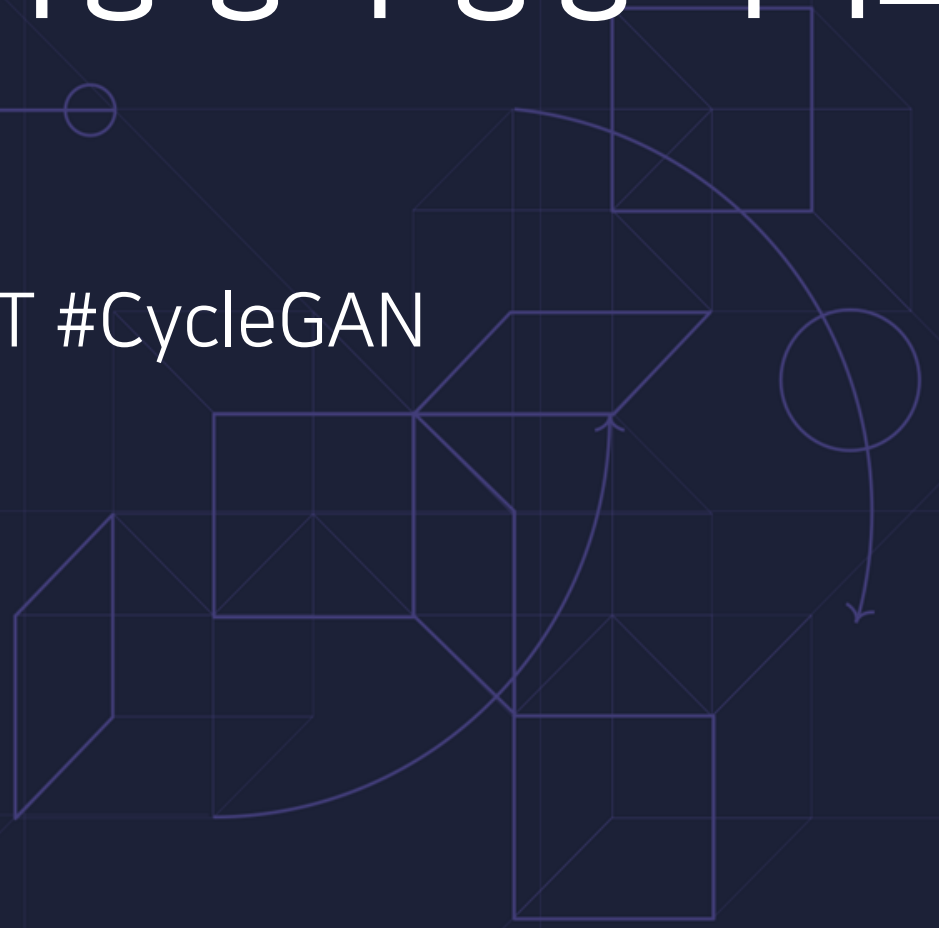
GPT GAN을 사용한 참여형 동화 생성 서비스 GATHERBOOK

#동화 #Text Generation #KoGPT #CycleGAN

NLP-06
(자,연어 한 접시)



boostcamp ai tech

About Team

Level1부터 함께 하고 있는 팀 **자, 연어 한 접시**를 소개합니다

이도훈

텍스트 데이터 수집 및 전처리
이미지 데이터 수집
문장 생성 모델 베이스라인
프로토타입 구축
프론트엔드 구축
동화 생성 모델 학습

김선재

텍스트 데이터 수집 및 전처리
이미지 데이터 수집
문장 추천 모델 실험
백엔드 구축
프로토타입 테스트 설문
동화 생성 모델 학습

차경민

텍스트 데이터 수집 및 전처리
이미지 데이터 수집
이미지 style-transfer 모델 학습
동화 style 이미지 증강
동화 생성 모델 학습

강진희

텍스트 데이터 수집 및 전처리
이미지 데이터 수집
키워드 추출 실험
프로토타입 테스트 설문
동화 생성 모델 학습

김태훈

텍스트 데이터 수집 및 전처리
이미지 데이터 수집
서비스 아키텍처 구성
프로토타입 테스트 설문
프로젝트 관리

목차

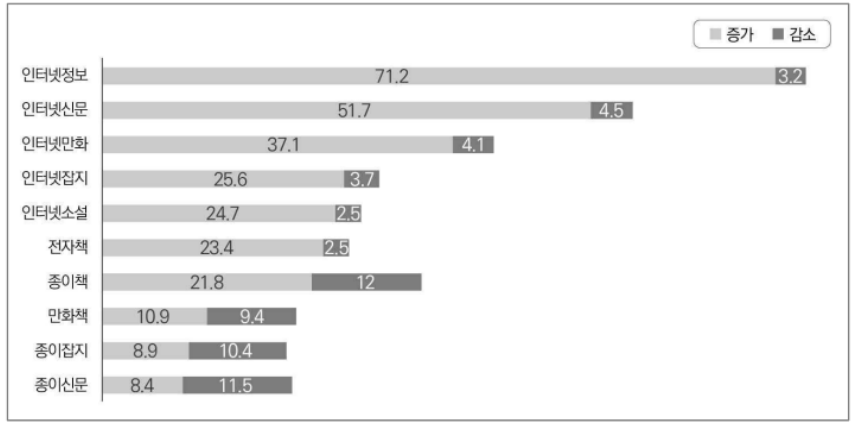
1. Intro

1-1. 프로젝트 배경

1-2. 서비스 소개

프로젝트 배경

- 기획 배경



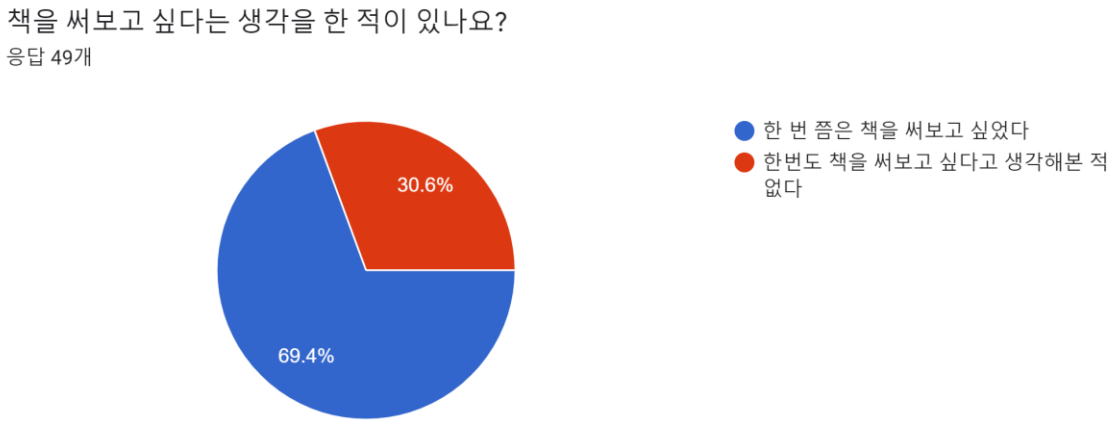
[그림 2-7] 코로나19 확산 이후 읽기 매체의 변화

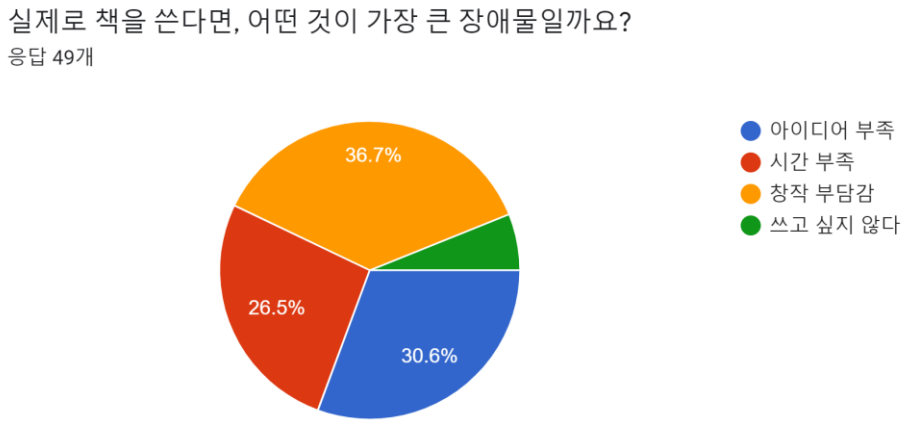
- 코로나19 확산 이후 디지털 기반의 콘텐츠 소비가 확대
- 전자책 시장의 성장과 나만의 전자책 만들기 콘텐츠의 인기
- 디지털 콘텐츠 시장의 성장에 따라 웹 기반 서비스 기획
- 아이부터 어른까지 전 연령대가 접근할 수 있는 장르적 특성에 따라 동화 서비스 기획

스마트미디어를 활용한 독서 생활화 방안 연구, 한국출판문화산업진흥원, 2022

프로젝트 배경

- 창작의 주요 걸림돌은 창작 부담감과 아이디어, 시간



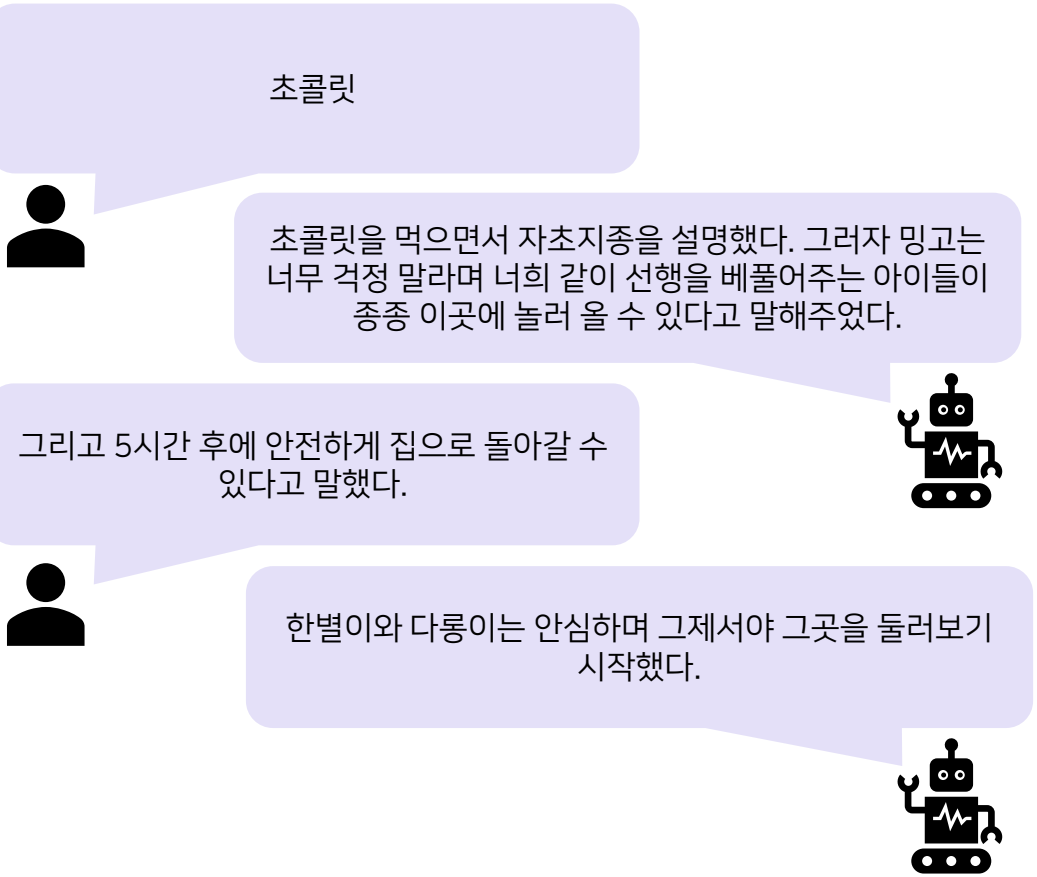


*10대 이하~50대 49명 대상 온오프라인 설문 조사 (2022.06.01-2022.06.05)

글쓰기를 돕는 서비스가 있다면?

서비스 소개

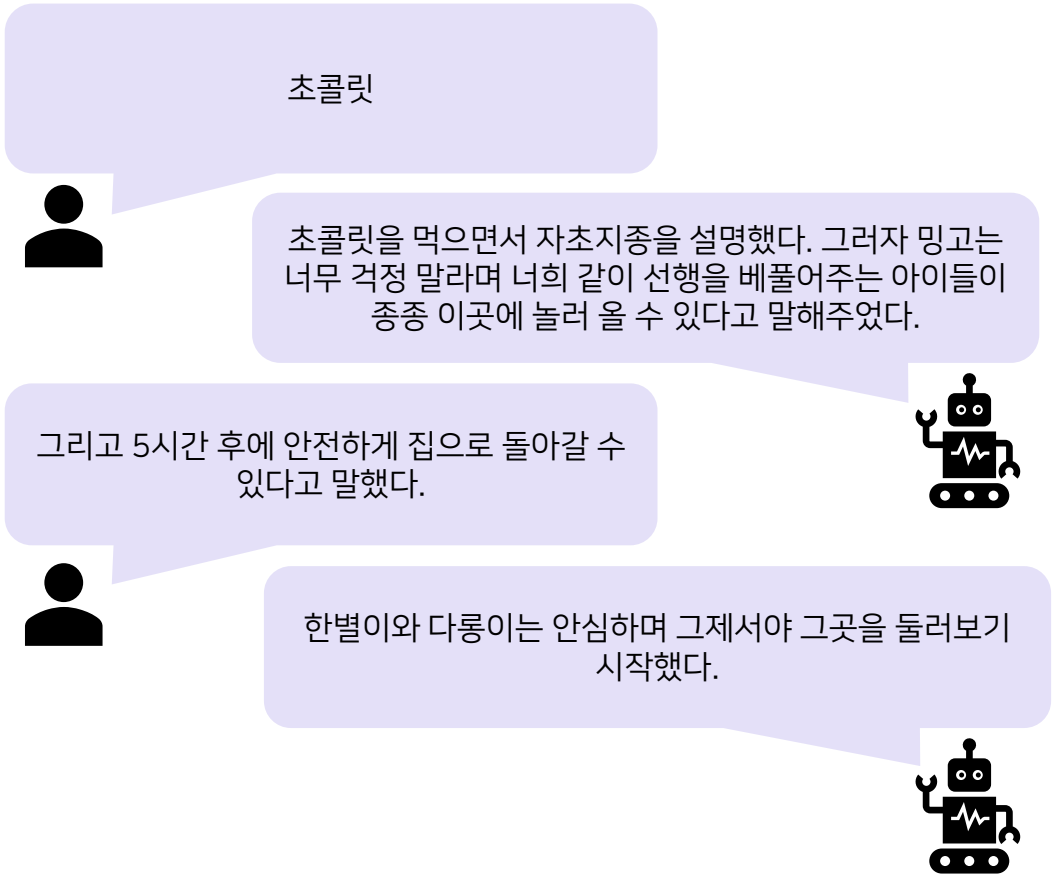
GATHERBOOK 은 사용자와 AI 모델이 번갈아 가며 동화를 창작하는 서비스 입니다



- GatherBook으로 나만의 스토리를 만들어보자
  - AI와 함께 동화를 작성하고
  - 동화에 어울리는 나만의 사진을 넣어
  - 나만의 동화 한 편을 완성한다
  - 작성한 글을 SNS 상에 공유해 나만의 스토리를 선보이자

서비스 소개

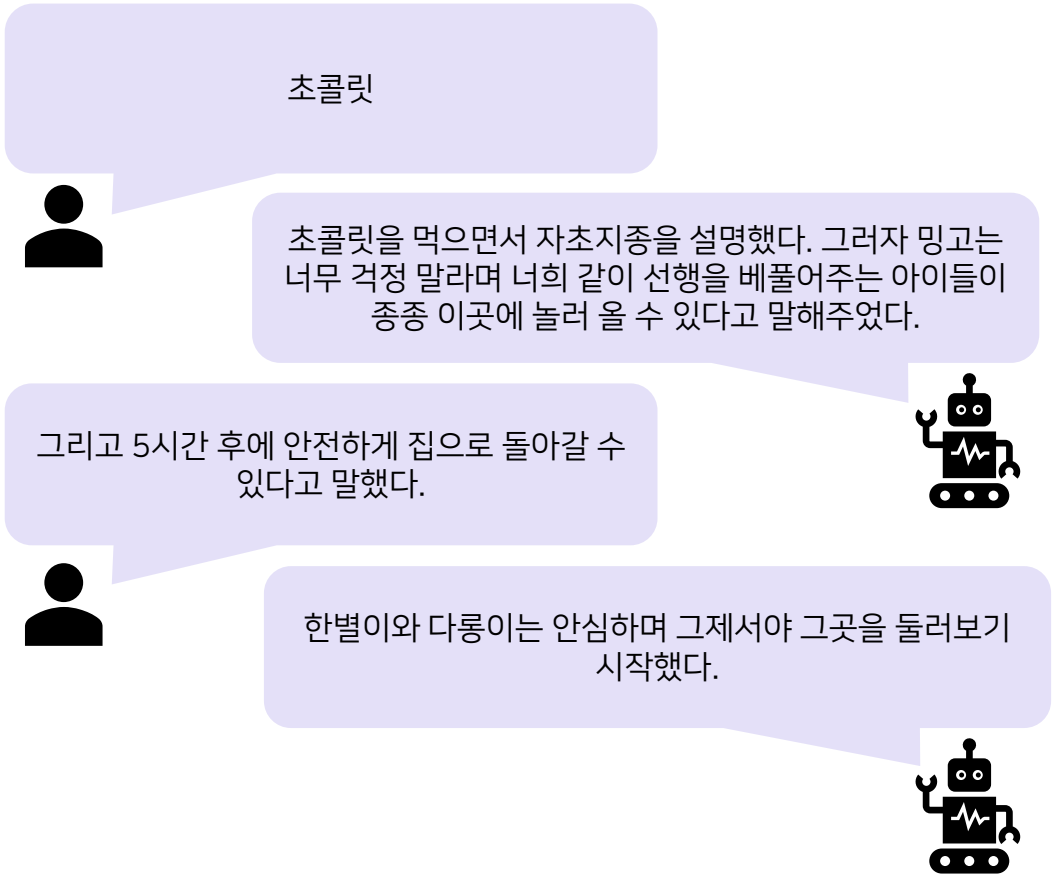
GATHERBOOK 은 재미있는 글쓰기를 지향하는 서비스 입니다



- GatherBook으로 창작의 부담에서 벗어나 쉽고 **재미있는 글쓰기**를 경험할 수 있습니다
  - GatherBook으로 아이와 함께 동화 만들기
  - GatherBook으로 나만의 동화 만들기
  - GatherBook 동화로 SNS에서 나를 표현하기

서비스 소개

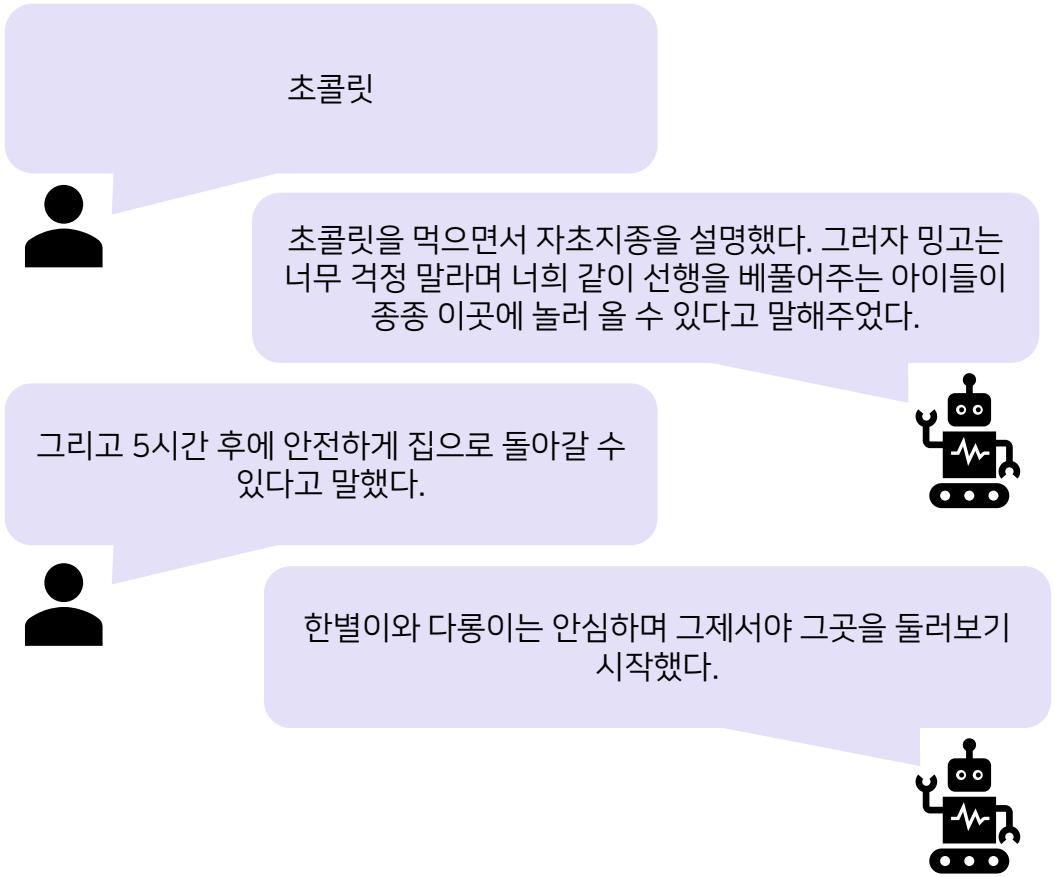
GATHERBOOK 은 쉽고 빠른 글쓰기를 가능하게 합니다



- GatherBook으로 동화 창작의 **아이디어**를 얻을 수 있습니다
  - AI가 제안하는 색다른 내용의 동화 전개
- GatherBook으로 창작의 **시간**을 단축할 수 있습니다
  - 다음 문장이 막막할 땐? GatherBook의 추천으로 이야기를 이어보자

서비스 소개

GATHERBOOK 은 쉽고 빠른 글쓰기를 가능하게 합니다



- GatherBook으로 동화 창작의 **아이디어**를 얻을 수 있습니다
  - AI가 제안하는 색다른 내용의 동화 전개
- GatherBook으로 창작의 **시간**을 단축할 수 있습니다
  - 다음 문장이 막막할 땐? GatherBook의 추천으로 이야기를 이어보자

2. Model

2-1. 데이터

2-2. 모델

데이터

- 동화 텍스트 데이터 수집

전래동화	세계 명작 동화	현대 창작 동화
청와대 어린이 전래동화 100편	그림형제 동화집 번역본 210편 이솝우화 동화집 번역본 222편	국립국어원 비출판물 말뭉치 444편 *신춘문예 동화 당선작 35편

- 텍스트 데이터 전처리
  - 작가의 주석, url, 괄호 안 부연설명, 특수기호 등 불필요한 문자 제거
  - 생성모델 학습 방식에 따른 데이터 구성 실험
    - 모든 동화 데이터가 병합된 전체 corpus의 문장 단위 학습
      →서로 다른 문서 간의 연관성이 반영되어 오히려 전체 맥락 단위의 부자연성 발생
    - 문서 단위의 batch 구성 후 문장 학습
      →생성 문장의 전체적 연결성 개선

*강원일보 등 15개 신문사 동화부문 당선작

데이터

- 동화풍 style Image 수집
  - 저작권 free image 사이트 pixabay, pinterest, Adobe Stock 등
  - 동화풍 이미지 149장 수집
  - CycleGAN 공식 학습 이미지 apple2orange, summer2winter, horse2zebra 등 논문 기재된 데이터셋 3.2GB

- VGG 사전학습 모델 활용 Style Image 증강

Content Image

+

Style Image

=

Result(=new Style Image)

모델 – 동화 생성 모델

- GPT-3 기반의 한국어 생성 모델

  Huggingface에 오픈소스로 공개되어 있는 GPT 기반 한국어 모델 비교

	Kakaobrain/KoGPT	SKT/KoGPT-trinity	SKT/KoGPT2
	GPT-3	GPT-3	GPT-2
parameters	약 62억 개	약 12억 개	약 1억 2500만개
layers	28개	24개	12개

  - Kakaobrain의 KoGPT는 Out-Of-Memory 문제로 인해 학습시키기 어려움
  - SKT의 KoGPT-trinity는 GPT-2 기반의 KoGPT2보다 성능이 좋다
  - Out-Of-Memory 문제가 발생하지 않는 선에서 가장 좋은 성능을 보이는 SKT의 KoGPT-trinity 사용

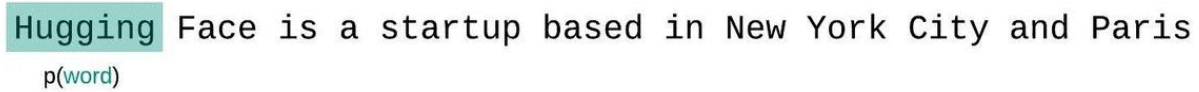
- Perplexity(PPL)
  - 언어 모델의 평가 지표

$$PPL(W) = \sqrt[N]{\frac{1}{P(w_1, w_2, w_3, \ldots, w_N)}} = \sqrt[N]{\frac{1}{\prod_{i=1}^{N} P(w_i | w_1, w_2, \ldots, w_{i-1})}}$$

Kichang Yang,"Transformer-based Korean Pretrained Language Models: A Survey on Three Years of Progress", https://arxiv.org/pdf/2112.03014

모델 – 동화 생성 모델 평가

- Perplexity 기반 모델 & 데이터 선택
  - Fixed length t = 1024



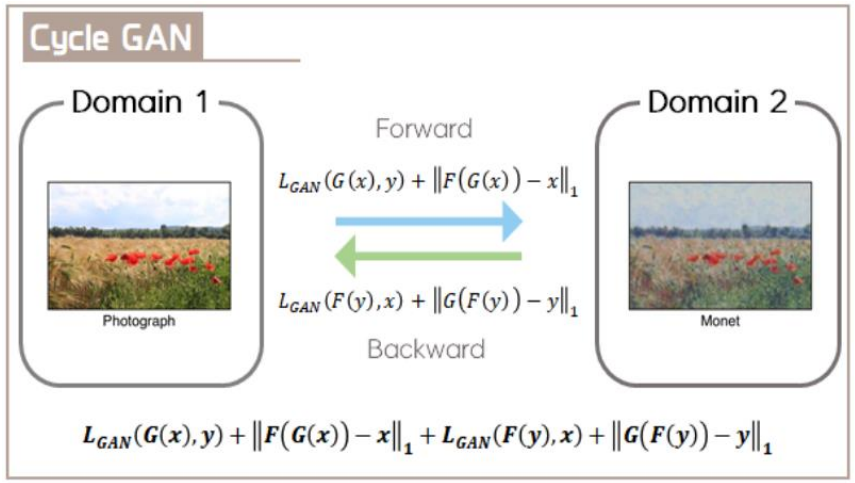
input	KoGPT2	KoGPT2 + 데이터 전처리	KoGPT2 + 동화별 학습	KoGPT-trinity + 동화별 학습
맑은 하늘	5.20	9.08	7.06	2.44
옛날에 고양이 한 마리가 있었어요	6.26	13.01	12.04	1.97
같은 날, 같은 도시에서 태어난 두 아이. 한 아이는 불쌍하게 거지 집에서 태어난 경민이였고 한 아이는 궁전 왕자로 태어난 도훈이였지요	8.01	224.09	177.61	4.59

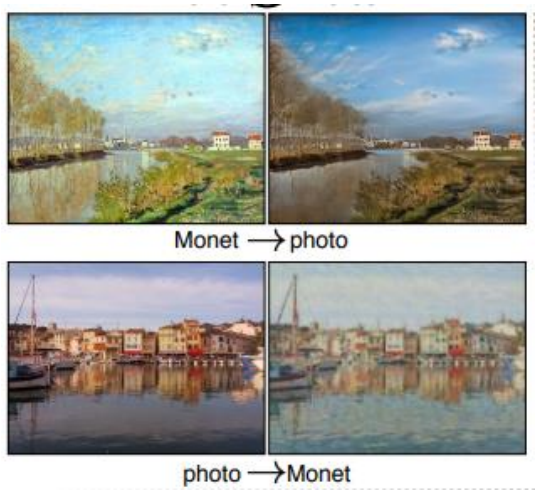
  - KoGPT2: inference 시 문장 잘림 현상
  - KoGPT2 & 데이터 전처리: 짧은 문장에서 perplexity값에 차이가 없지만, 긴 문장에 대한 perplexity값에 차이
  - KoGPT2 & 동화별 학습(문서 단위 batch): 문맥 이해 성능의 개선
  - KoGPT-trinity & 동화별 학습(문서 단위 batch): 긴 문장에 대한 출력 성능 개선 & perplexity 값의 개선

Calculating PPL with fixed-length models, Hugging Face, https://huggingface.co/docs/transformers/perplexity

모델 – 이미지 style-transfer 모델

- GAN 기반 style-transfer 모델
  - 사전학습 모델은 입력 받은 이미지에 대한 학습이 필요->inference 시간 소요
  - Inference 과정에서 generator를 활용한 GAN 모델 탐색





- CycleGAN
  - Unpaired Image-to-Image translation 가능
    - 일반적인 GAN 모델은 paired dataset이 필요하다. 그에 따른 추가 이미지 수집이 필요
    - Unpaired image dataset->동화풍 이미지 수집 만으로 학습이 가능하다. 추가 데이터 수집이 필요 없음
    - 사전학습 모델을 통해 동화풍 이미지 데이터를 증강하고 활용할 수 있다.

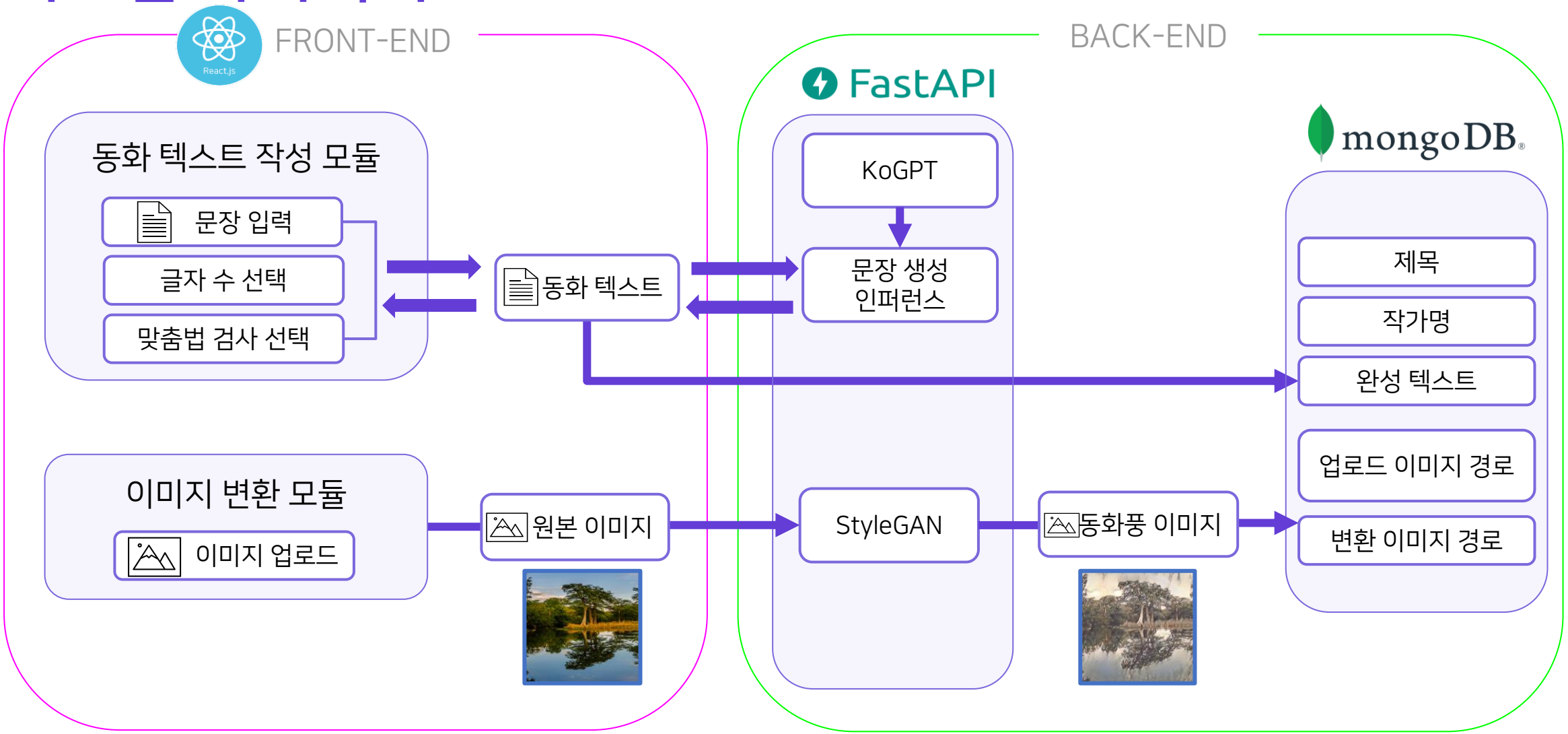
Jun-Yan Zhu et al. , "Unpaired Image-to-Image Translation using Cycle-Consistent Adversarial Networks" , https://arxiv.org/pdf/1703.10593

3. Product Serving

3-1. 시스템 아키텍처

3-2. 프로토타입 테스트

시스템 아키텍처


FRONT-END
BACK-END
React.js
FastAPI
mongoDB
동화 텍스트 작성 모듈
문장 입력
글자 수 선택
맞춤법 검사 선택
동화 텍스트
KoGPT
문장 생성
인퍼런스
제목
작가명
완성 텍스트
이미지 변환 모듈
이미지 업로드
원본 이미지
StyleGAN
동화풍 이미지
업로드 이미지 경로
변환 이미지 경로

프로토타입 테스트

Streamlit 기반의 프로토타입 구축

사용자의 입력

옛날 옛날에 아주 콧대 센 공주님이 한 분 살고 계셨어요.

사용자 입력에 이어 모델이 생성한 내용

그녀는 구혼자들이 나타나면 몇 가지 수수께끼를 내서 남자가 못 알아듣게 했어요. 그녀는 누구든 수수께끼를 하나라도 맞힐 수 있다면 자신의 남편으로 삼을 용의가 있다고 말했답니다.

많은 사람들이 찾아왔지만, 그 누구도 수수께끼를 풀지 못했어요.

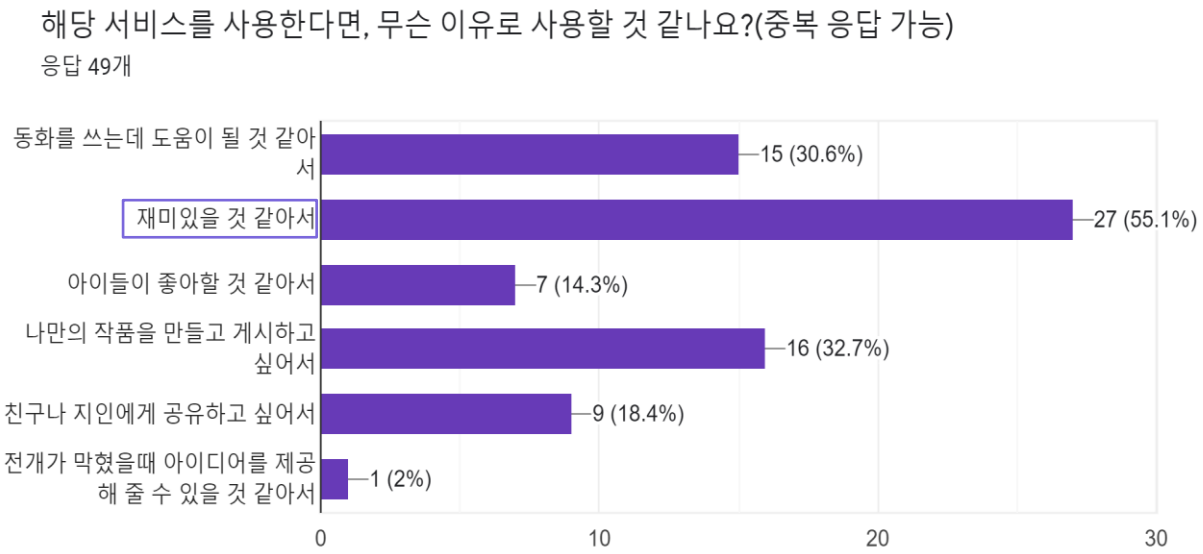
그러자 보다 못한 누군가가 왕에게 얘기 했어요.

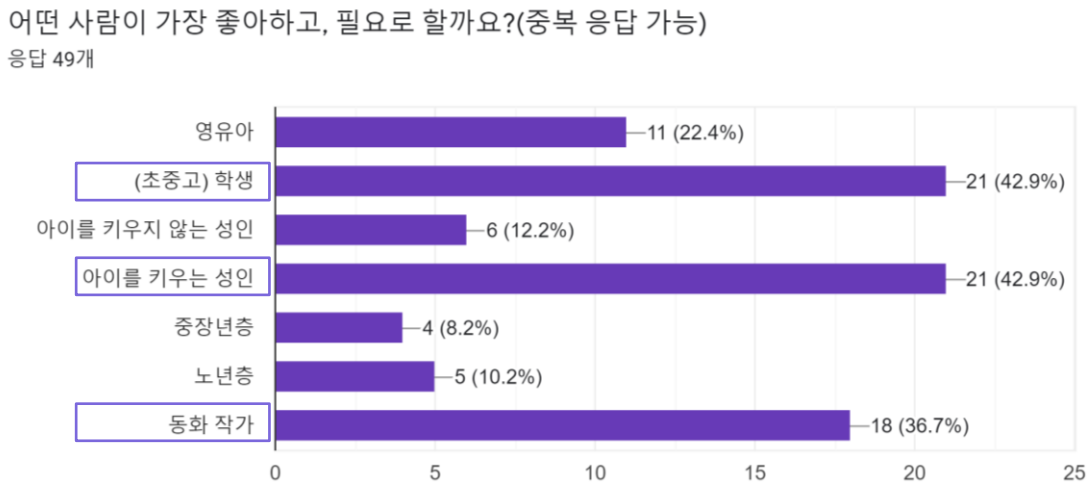
프로토타입 테스트

10대 ~ 50대 49명에게 *온오프라인 설문으로 서비스에 대한 의견을 구했습니다.

(*오프라인: 2022 서울국제도서전 2022.06.01~2022.06.05 온라인: 구글 설문)

서비스 선호 조사





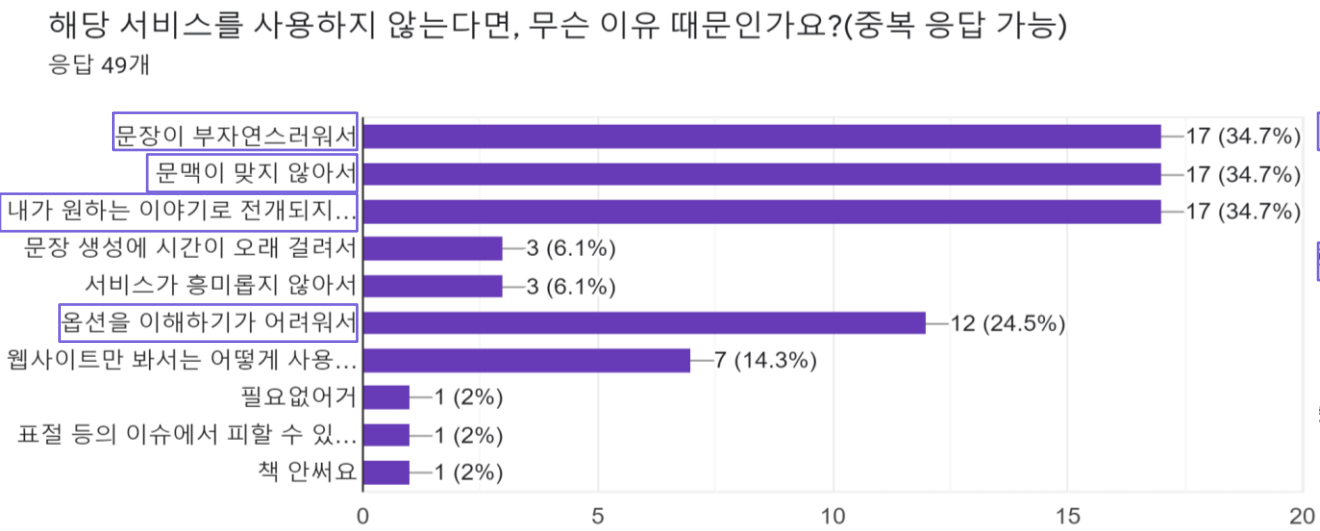
동화 작가를 비롯한 다양한 계층이 재미있게 사용할 수 있는 서비스

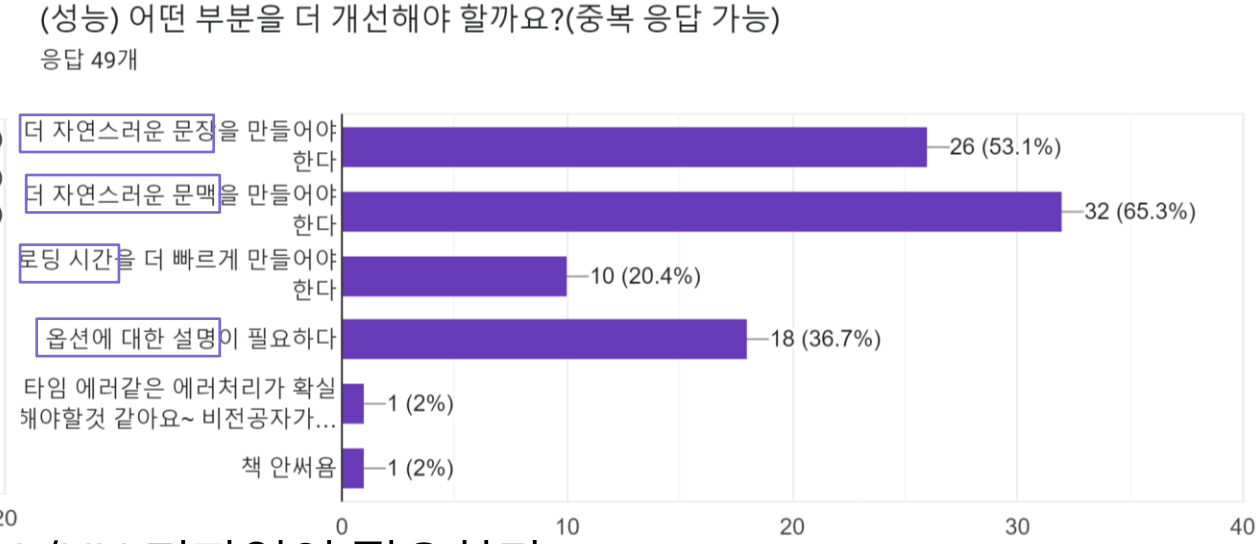
프로토타입 테스트

10대 ~ 50대 49명에게 *온오프라인 설문으로 서비스에 대한 의견을 구했습니다.

(*오프라인: 2022 서울국제도서전 2022.06.01~2022.06.05 온라인: 구글 설문)

서비스 개선 사항





문장 생성 성능 개선과 로딩 시간 개선, 이해하기 쉬운 UI/UX 디자인이 필요하다

4. Conclusion

4-1. 시연 영상

4-2. 한계점 & 후속 개발

시연 영상

http://gather-book-front.herokuapp.com/

GatherBook 동화 생성 예시

맑은 하늘과 뭉게구름이 산 위로 모여드는 것 같아.

걸음을 옮길 때마다 고슴도치 한스는 자신의 아내에게 말했지요. “저기를 봐!” 라며 늑대가 말했어요.

이렇게 아름다운 새가 왜 이 나무 위에 앉아 있는 거지?

여우는 계속 숲으로 들어가며 생각했어요.

옛날에 호랑이 한 마리가 살고 있었어요.

그런데 이 호랑이는 바로 옛날 옛적에 아주 멀리 떨어진 궁전에서 몰래 생활을 하고 있었던 거예요.

호랑이님, 여기 계셨군요! 그래, 어디를 그리 바삐 가시나요?

궁전 구석구석을 돌아다녀보아도 제게 발각되지 않는 게 없어. 그 말을 들은 고양이는 냉큼 가서 지갑을 들고 나왔답니다.

한계점 & 후속 개발

1. 등장인물의 비일관성

입력 : "해와 달이 된 오누이"

출력 : 해와 달이 된 오누이 둘을 보며 할머니는 웃었어요. 그리고 달을 보면서 할머니의 눈물을 훔쳐도 보았고요. 어느 날 할머니가 손을 쥐어짜며 말했어요. 이번에는 개똥이가 한 번 읽어보아라!

- 새로운 인물의 등장: 출력 문장에서 할머니, 개똥이 등장
- 이야기 주체의 변화: 생성된 문장에서 할머니가 등장하면서 이야기의 중심이 할머니로 옮겨짐
- 사용자가 기대하는 흐름과 전혀 다른 등장인물(예:개똥이)이 나타날 수 있다

한계점 & 후속 개발

1. 등장인물의 비일관성 ➔ 등장인물의 고정
   - 사용자로부터 등장인물 설정을 입력 받고, 등장인물을 고정하는 문장 생성 방식
   - 학습 데이터에 동화(문서) 단위로 모든 등장인물을 [PER1], [PER2] 등 마스킹 처리

     → 이야기의 등장인물을 고정하는 효과
   - 사용자의 등장인물 입력과 [PER1] 토큰의 매칭

     → 사용자가 원하는 등장인물의 사용

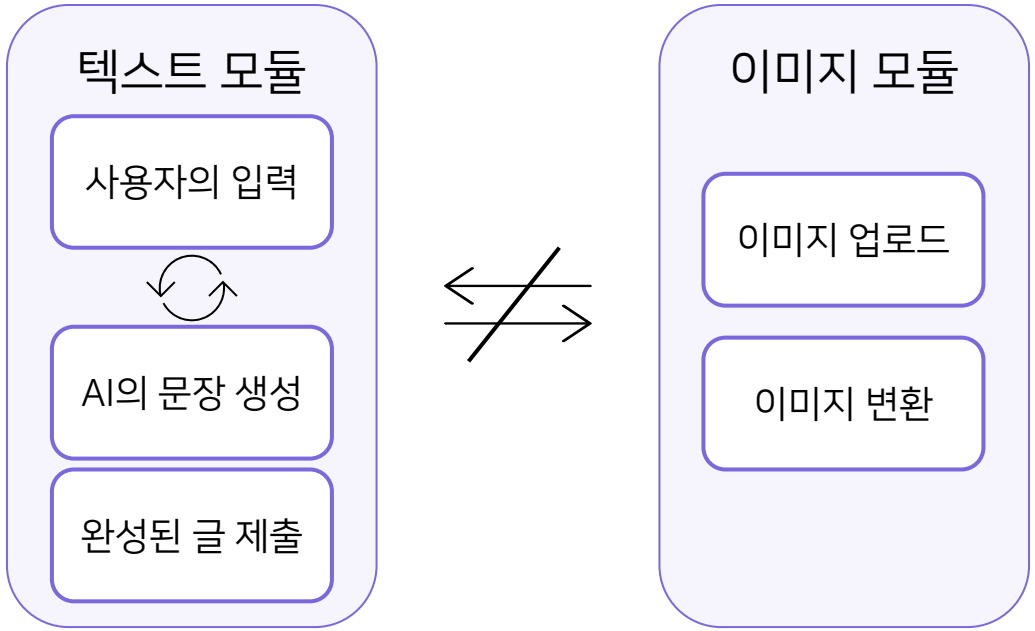
입력 : "어느 날 [주인공]가 집을 나섰습니다." + (주인공: 공주, 등장인물: 자매, 왕자)

출력 : 어느 날 [주인공]가 집을 나섰습니다. 그러자 두 [PER1]들이 냉큼 그녀에게 달려들어 용암처럼 그녀들을 삼켜버렸어요. 하지만 [주인공]가 구한 [PER1]은 왕자님 덕분에 가까스로 화를 누그러뜨릴 수 있었답니다. 다음날 [PER2]가 [주인공]에게 가 자신이 [주인공]의 약혼자라 말하며 자신의 황금 머리카락 세 개를 제시했어요.

한계점 & 후속 개발

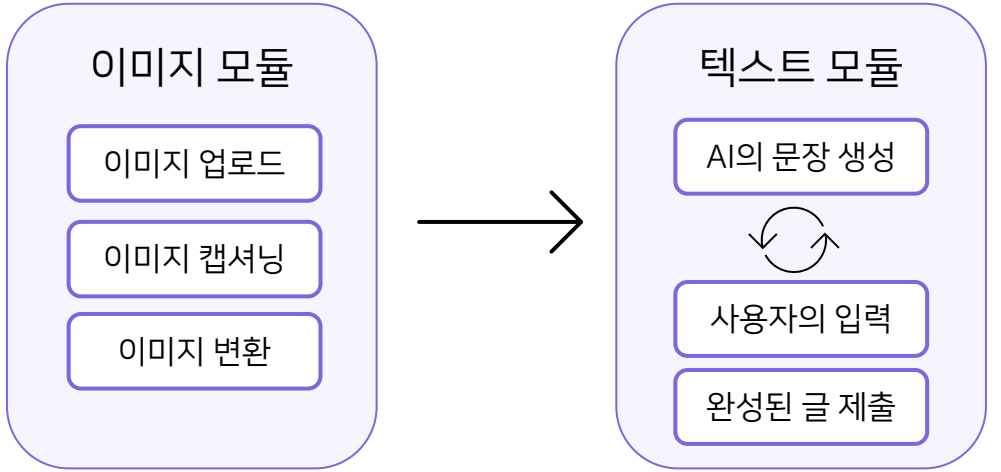
2. 이미지와 글의 연관성 개선

- 사용자가 완성된 이야기와 어울리는 이미지를 가지고 있어야 한다
  - 사용자와 AI가 번갈아가며 창작하는 방식 → 완성된 글의 흐름을 미리 알기 어렵다
- 동화풍 style-transfer는 이미지와 이야기의 주제를 연결시키기 어렵다



한계점 & 후속 개발

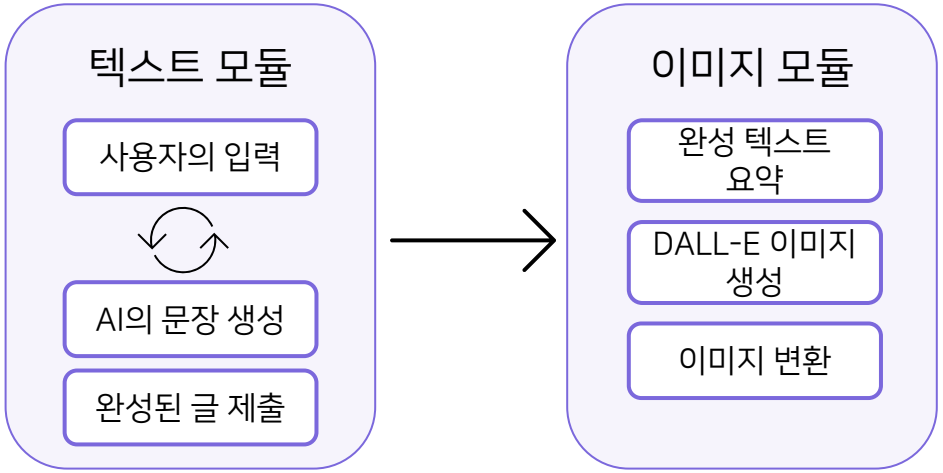
2. 이미지와 글의 연관성 개선 ➔ 이미지 캡셔닝 활용
   - 사용자가 입력한 이미지에 어울리는 글을 창작하는 서비스
   - 이미지를 보고 동화의 첫 문장을 AI가 작성하는 기능
   - AI Hub의 한국어 이미지 설명 데이터셋(MS COCO 한국어 번역본) 기반 이미지 캡셔닝 학습
   - 캡셔닝 결과를 동화식 문체(해요체)로 변환



Oriol Vinyals et al. , "Show and Tell: A Neural Image Caption Generator", https://arxiv.org/pdf/1411.4555.pdf

한계점 & 후속 개발

2. 이미지와 글의 연관성 개선 ➔ DALL-E 활용한 삽화(표지) 생성
   - 사용자의 문장 입력과 모델의 문장 생성으로 동화 완성
   - 완성된 텍스트의 retrieval 요약
   - DALL-E 입력 후 요약에 맞는 이미지 생성
   - Inference 시간 단축이 중요



https://openai.com/blog/dall-e/ , Aditya Ramesh et al., "Zero-Shot Text-to-Image Generation", https://arxiv.org/pdf/2102.12092

한계점 & 후속 개발

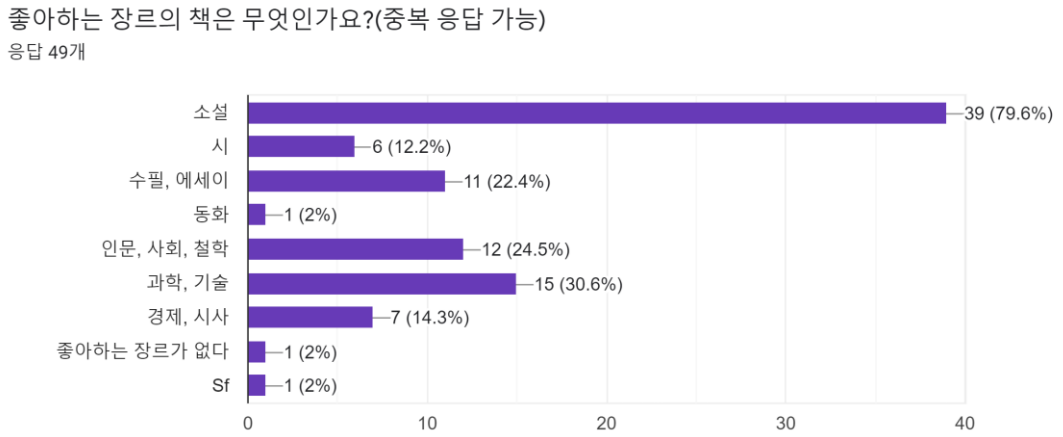
3. 장르의 확장

- 프로토타입 설문 결과, 다양한 장르 선택에 대한 의견

장르 선택

장르선택 > 키워드(n개 선택) 등

동화말고도 소설이나 다른 책 유형을 지원하는 서비스도 있었으면 좋겠습니다



- 동화 외 선호도가 높은 단편 소설, 수필, 시 장르의 데이터를 추가로 수집하여 학습시키고 장르 선택 기능을 추가하여 사용자에게 장르 선택의 폭을 넓힐 수 있다

한계점 & 후속 개발

4. 완성 동화를 게시 전 표현의 내부 필터링 추가
   - 어떤 입력이 들어와도 웹페이지 게시판에 게시 가능
   - 서비스 관점에서 부적절한 게시글에 대한 필터링이 필요
   - 비속어, 혐오 표현이 포함된 글에 대한 필터링
   - 우울감을 나타내는 글에 대한 필터링
   - 정치적, 차별적 표현이 포함된 글에 대한 필터링

     ➔ 딥러닝 기반 혐오 표현 필터링 모델 구축

한계점

5. 출간용 서비스를 위한 창작 전문성 – 표절 방지 및 저작권 문제

- 동화 작가의 창작 보조 도구로 사용할 수 있을까?
  - 완전한 출판물의 보조 도구로 사용되려면 표절 문제를 해결할 수 있어야 한다
  - 출간된 창작물에 대한 표절 검사 기능이 필요
  - 저작자의 동의를 받은 출판물의 본문 데이터 확보가 필요하다
- AI 서비스와 공동 작업한 창작물에 대해 저작권을 인정받을 수 있어야 한다

저작권법 제2조 제1호는 '인간의 사상이나 감정을 표현한 창작물'을 저작물로 규정하고 제2호는 '저작물을 창작한 자'를 저작자로 규정하고 있으며 제10조 제1항은 저작자는 저작인격권 및 저작재산권을 가진다고 규정하고 있다. 저작권법 규정에 따라 인간의 사상이나 감정을 표현한 창작물을 창작한 '자'는 저작권을 향유할 수 있게 된다. 저작권 보호의 목적을 창작자의 지적 활동의 결과물에 대한 보상이라고 보는 측면과 현행 저작권법상 저작물이 인간의 사상이나 감정의 창작적 표현이라고 하는 개념 정의에 충실하면 저작자는 자연인에 한정된다고 해석될 수 있고, 자연인이 아닌 인공지능은 저작자로 볼 수 없게 된다.34) 현행 저작권법 규정에 따라 자연인만이 저작권자가 될 수 있다고 해석하는 것이 합리적이라고 본다.

레퍼런스

- 한국출판문화산업진흥원, 스마트미디어를 활용한 독서 생활화 방안 연구
- Jun-Yan Zhu et al. , “Unpaired Image-to-Image Translation using Cycle-Consistent Adversarial Networks” , https://arxiv.org/pdf/1703.10593
- Kichang Yang,“Transformer-based Korean Pretrained Language Models: A Survey on Three Years of Progress”, https://arxiv.org/pdf/2112.03014
- [CycleGAN] Unpaired Image-to-Image Translation using Cycle-Consistent Adversarial Networks 논문 리뷰, https://di-bigdata-study.tistory.com/9
- Oriol Vinyals et al. , “Show and Tell: A Neural Image Caption Generator”, https://arxiv.org/pdf/1411.4555.pdf
- Aditya Ramesh et al., “Zero-Shot Text-to-Image Generation”, https://arxiv.org/pdf/2102.12092
- https://openai.com/blog/dall-e/
- [CycleGAN 공식 train datsets] https://people.eecs.berkeley.edu/~taesung_park/CycleGAN/datasets/

감사합니다

boostcamp ai tech

5. Appendix

5-1. 발생했던 이슈

5-2. 개발 과정의 실험

5-3. 프로젝트 관리

발생했던 이슈

- 특정 키워드(단어)에 대한 일관된 문장 생성

입력 : "맑은 하늘 "

출력 : 맑은 하늘, 푸른 색을 감싸 안은 모래사장의 모습에 막연한 동경을 품게 된 것이다. 그러므로 이렇게 기도하여라, … 하지만 난 아직도 그날 눈이 펑펑 내리던 날을 기억하고 있다.

입력 : "맑은 하늘 "

출력 : 맑은 하늘, 푸른 색을 감싸 안은 모래사장의 모습에 막연한 동경을 품게 된 것이다. 그러므로 이렇게 기도하여라, … 그러나 나는 확실하다. 난 아직까지도 하늘의 별들을 헤아릴 수 있고, …

  - 특정 키워드 맑은 하늘에서 일관된 문장 생성 (다른 키워드는 발생하지 않음)
  - 맑은 하늘 외 일관된 문장 생성 이슈가 발생하는 키워드 파악 필요

- Inference 시 repetition_penalty 값을 높임으로써 해결 가능
  - 사용자에게 repetition_penalty = 1.5 추천(default 설정)

발생했던 이슈

- 단어 1개만 입력 시 출력이 고정되는 이슈

> 입력 : "여름"
>
> 출력 : 여름이 되어 야외에서 더 이상 먹을 걸 찾을 수 없게 되자, 쥐는 돼지비계 항아리를 생각해내곤 말했어요. 가자 고양이야, 이런 때를 대비해서 돼지 비계가 든 항아리만을 보관해두었잖니…

  - 하나의 단어(예: 여름)만 입력하면 동일한 출력이 반복된다
    - GPT 모델의 구조: 이전 토큰의 시퀀스 기반으로 다음 토큰을 추론
    - 하나의 단어(=하나의 토큰)만으로 다음 토큰을 추론할 시, 결과가 동일하게 고정된다
  - 하나의 토큰에 대한 여러가지 결과 출력이 필요

- 해결 방안
  - 스페셜 토큰을 추가하여 출력 케이스의 다양성을 높인다 → 시간 부족으로 인해 시도하지 못함
  - 사용자에게 2개 이상의 단어 입력을 권장 → 사용자 입력 단계의 rule 추가

발생했던 이슈

- 텍스트 생성 시 대명사/지칭어 발생 문제

  - 사용자와 번갈아가며 작성하는 컨셉이기 때문에, 모델이 생성한 문장을 사용자가 이해할 수 있어야 다음 문장을 창작할 수 있다.
  - 사람, 사물 등 여러 객체가 등장할 경우, 대명사의 사용은 사용자의 이해를 방해할 수 있다.

입력 : "같은 날, 같은 도시에서 태어난 두 아이. 한 아이는 불쌍하게 거지 집에서 태어난 톰이였고,
한 아이는 궁전 왕자로 태어난 에드워드였지요."

출력 : 그녀는 왕이 될 자격이 없어 보였기 때문이지요.

  - 출력의 그녀가 입력의 톰을 가리키는지, 에드워드를 가리키는지 알기 어렵다.

- 학습 데이터의 모든 등장인물을 [PER1], [PER2] 과 같은 스페셜 토큰으로 마스킹하고, 등장인물 표현을 고정한다. (후속개발 1번과 동일)

발생했던 이슈

- 문장 종결 어미의 통일

입력 : "구름을 헤치며 날아가고 싶다."

출력 : 구름을 헤치며 날아가고 싶다. 하지만 그는 소망만 하고 있었죠.

입력 : "구름을 헤치며 날아가고 싶어."

출력 : 구름을 헤치며 날아가고 싶어. 그런 다음 그가 긴 의자 위에 팔다리를 쭉 뻗으며 누웠어요..

  - 입력 문장의 종결 어미에 관계 없이 해요체(~요)로 끝나는 경향이 있다.
    - 학습 데이터의 해요체 비중이 높기 때문인 것으로 파악
    - 입력 문장의 길이가 길면(2문장 이상) 입력 문장의 종결 어미를 따른다. 길이가 짧은 입력에서 발생하는 문제

- 입력 문장의 종결 어미를 분리하고 해요체로 수정하여 모든 문장을 동화풍의 해요체로 변경 시도
  - 종결어미 분리 후 해요체 합성 시 맞춤법 교정→inference 시간 소요 문제
  - 종결어미 분리는 한국어 토크나이저 성능에 의존하는 경향이 높으며 완전하지 않다.

- 스페셜 토큰 추가 및 동화 문서 단위의 배치 작업 후 문장 생성 성능의 향상→개선

발생했던 이슈

- 마침표 유무에 따른 문장 생성의 차이

입력 : "구름을 헤치며 날아가고 싶어"

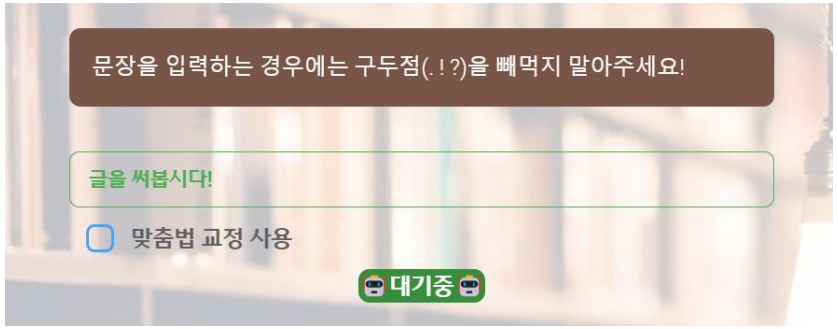
출력 : 구름을 헤치며 날아가고 싶어 했어요. 하지만 그는 붙잡지 못했지요. …

입력 : "구름을 헤치며 날아가고 싶다"

출력 : 구름을 헤치며 날아가고 싶다 말했어요. 하지만 그는 소망만 하고 있었죠. 왜냐하면 …

  - 입력 문장에 마침표(구두점 .?!)이 없는 경우, 문장이 끝나지 않았다고 인식하여 종결 어미부터 생성
  - 마침표가 없는 문장을 사용자가 의도한 하나의 완결된 문장으로 볼 것인지, 오타로 처리할지 규정할 필요
  - 사용자의 의도로 일부러 마침표 없이 완결된 하나의 문장일 수도 있다. 이것을 모델이 판단하기에는 어려움

- 사용자에게 입력 시 구두점을 넣어 달라는 안내 문구 추가



발생했던 이슈

- 이미지 inference 시간
  - VGG 기반 사전학습 모델은 결과가 잘 나오지만, inference 시간이 오래 걸린다

- 해결 아이디어
  - 저화질 저성능의 가벼운 모델과 학습 모델 2가지 버전 사용
    → 사용자에게 저성능 버전 preview
    → 로딩 시간의 지루함 해결
  - GAN 모델 아키텍처 탐색

- 서비스의 중심 point 는 생성 파트에 집중
  → 이미지 변환의 로딩 시간 개선이 필요
  → GAN 모델 탐색

개발 과정의 실험

Style-transfer 모델 실험

1. VGG 기반 Pre-trained Model 실험
   - 일반적인 사진이 동화풍 스타일 이미지와 합쳐지며 괜찮은 결과가 생성
   - 하지만 inference 시간 소요가 문제
   - 이미지가 들어올 때마다 학습을 통해 output을 제공해야 하는 문제

2. GAN 계열 CycleGAN Model 실험
   - 사전에 학습을 진행하고 inference 과정에서 generator로 output을 만드는 GAN 모델 탐색
   - Pair-wised data 필요 x
     ➔ 데이터 수집&paired-labeling 이 따로 필요 없으며, 사전학습 모델 실험과 비슷한 결과를 보인다는 장점을 가진 CycleGAN 실험
   - 한정된 대회 기간 내에서 paired dataset을 만드는 것은 비효율적이라는 생각에 CycleGAN 선택

개발 과정의 실험

Style-transfer 모델 실험

VGG 기반 Pre-trained Model 실험 결과

GAN 계열 CycleGAN Model 실험

1) 일반 이미지 → 동화 이미지 변환

2) 동화 이미지 → 일반 이미지

개발 과정의 실험

문장 키워드 추출 실험

완성 글에서 키워드 추출 후 키워드에 알맞은 이미지를 검색API로 가져올 계획

- TextRank 알고리즘
  - 그래프 알고리즘 기반의 extractive 키워드 추출의 일종
  - Konlpy의 Okt(구 Twitter), Komoran, Hannanum 등 다양한 토크나이저 사용 가능
  - 1개 문서에서만 키워드 추출 → inference 시간 가장 짧은 장점
- TF-IDF
  - Extractive 키워드 추출 기법 중 대표적인 방법
  - 문서 내 자주 등장하는 단어 & 다른 문서에서는 등장 빈도가 낮은 unique를 고려하는 기법
  - 모든 문서에 대한 탐색 필요 → inference 시간 소요
- KeyBERT
  - BERT 기반의 키워드 추출 모델
  - 토크나이저 로딩, 텍스트 임베딩에서 시간이 많이 소요 되는 단점

개발 과정의 실험

문장 키워드 추출 실험

완성 글에서 키워드 추출 후 키워드에 알맞은 이미지를 검색API로 가져올 계획

- 예제 실험) 이무기에게 제물로 바쳐진 처녀를 구한 청년에 관한 백일홍 설화

	TextRank	TF-IDF	KeyBERT
상위 5개 키워드	처녀, 청년, 이무기, 그, 사람	마을, 청년이. 이무기, 처녀, 제사를	처녀, 비단, 꽃처럼, 신부, 이무기
Inference 시간	1.57초	22.14초	22.29초

  - TextRank의 inference 시간이 가장 빠르고, 적절한 키워드 추출 성능을 보인다

- 키워드 추출에 대한 검색 API의 이미지 매칭에서 *이슈 발생 ➔ 서비스에서 키워드 추출 단계 제외
  - *뉴스, 블로그, 카페 등 출처가 불분명한 이미지 및 부적합한 이미지 검색

개발 과정의 실험

문장 추천 모델 실험

- KNN 추천 알고리즘
  - 유클리드 거리 기반의 KNN 알고리즘으로 다음에 올 문장을 추천하는 로직 실험
  - 단어 단위로 벡터 표현할 경우, 유사도 높은 순으로 단어를 단순 배열하는 형태➔문법적 오류 발생
  - 유사도 기반 추천은 비슷한 문장을 추천하는 것. 다음에 이어질 문장을 추천하지 않는다
  - Sequential 이해 필요

- MCMC (Markov Chain Monte-Carlo)
  - Sequential 이해 가능
  - 생성모델처럼 새로운 샘플을 만들어서 데이터 증강 & 학습하는 방식
  - 새로운 샘플은 가지고 있는 데이터의 분포에서 크게 벗어나지 않게 생성된다
  - 문장 생성 모델(transformer기반 GPT 모델)의 성능이 기대 이하일 경우 사용
    - 딥러닝 기반의 GPT 사용을 우선

프로젝트 관리

- 실험 결과 & 아이디어 공유 ➔ Notion

📋 스크럼 보드

표에서 계획 중/해결 중/해결한 Issue를 작성해주세요

보드 보기 | 표 | + 보기 추가

시작 전 1	진행 중 6	완료 2
Docker 이도훈	동화풍 이미지 수집(추가) Jinhee Kang	문장단위 생성 이도훈 차경민
+ 새로 만들기	이미지 검색 api 탐색 Jinhee Kang	데이터 문장단위 전처리 이도훈 오리 차경민
	Quantization 이도훈	+ 새로 만들기
	Pruning	
	seq2seq 테스트 차경민	
	Perplexity Jinhee Kang	

👣 프로젝트 진행 리포트

표 | + 보기 추가

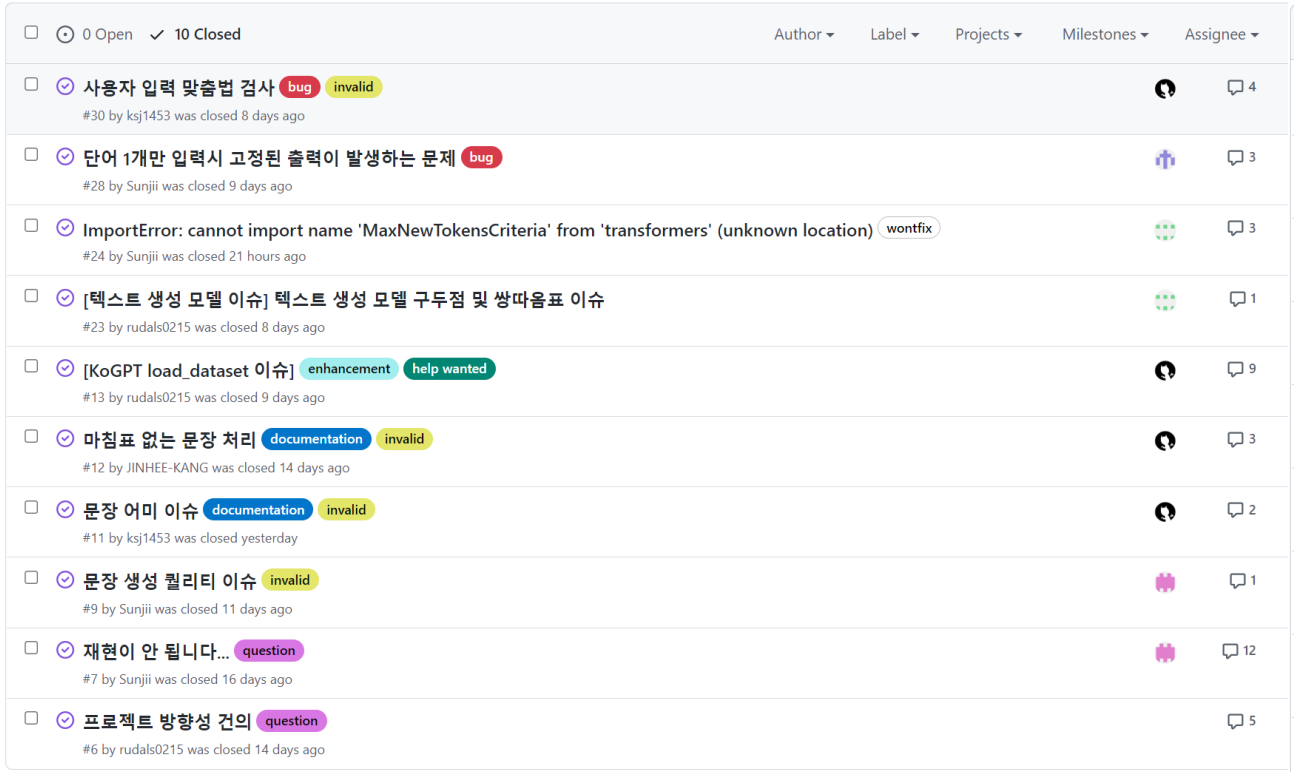
이름	Task	반영 일자	최종 업데이트
style transfer pretrained 베이스 모델 코드 실험	style-transfer	2022년 5월 15일	2022년 5월 18일 오후 4:38
KoGPT 모델 베이스 코드 실험	문장 생성	2022년 5월 16일	2022년 5월 16일 오후 10:43
mcmc, 맞춤법 검사 API	문장 추천	2022년 5월 16일	2022년 5월 16일 오후 10:43
키워드 추출(TextRank)	키워드 추출	2022년 5월 16일	2022년 5월 16일 오후 10:44
KoGPT 스페셜 토큰 추가	문장 생성	2022년 5월 16일	2022년 5월 16일 오후 11:13
cycle-gan 모델 실험	style-transfer	2022년 5월 17일 → 2022년 5월 19일	2022년 5월 23일 오후 3:19
cycle-gan inferece 코드 작성 및 실험	style-transfer	2022년 5월 23일	2022년 5월 24일 오전 9:20
KoGPT2 다양한 디코딩 방법 찾기	문장 생성		2022년 5월 24일 오전 3:14
프로토타입 테스트 케이스	문장 생성	2022년 5월 24일	2022년 5월 24일 오전 9:54
프로토타입 설문 결과	service	2022년 6월 1일	2022년 6월 2일 오전 10:12
개인 회고_T3007 강진희	회고		2022년 6월 11일 오후 4:25

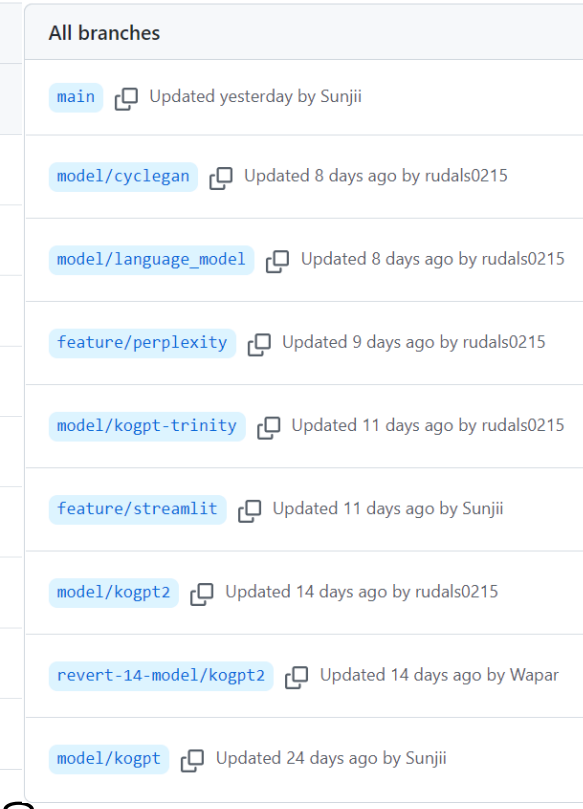
+ 새로 만들기

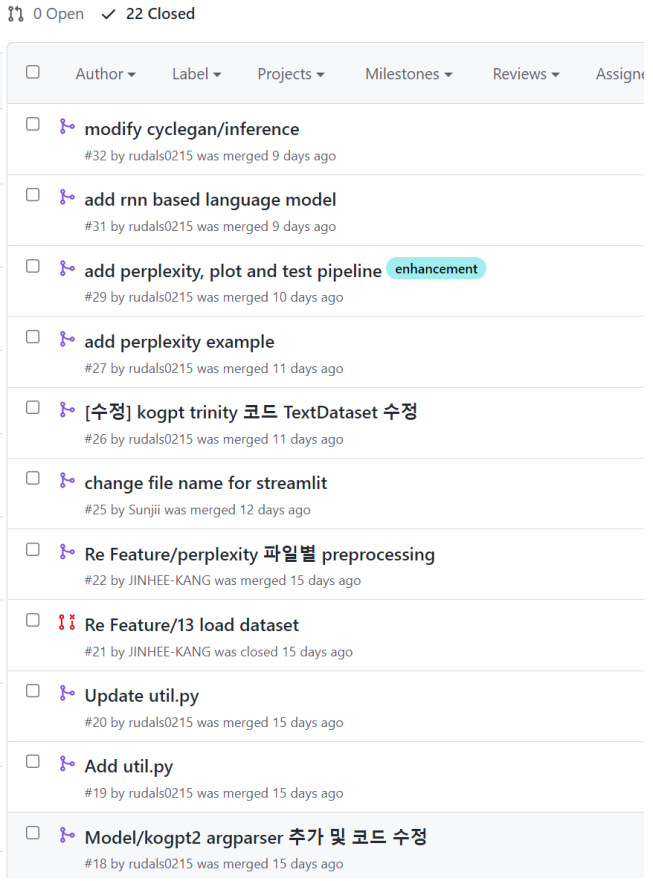
- 모델 실험 결과와 레퍼런스, 보고서 형식의 자료는 notion으로 작성 및 공유

프로젝트 관리

- 이슈 & 기능 관리 ➔ Github



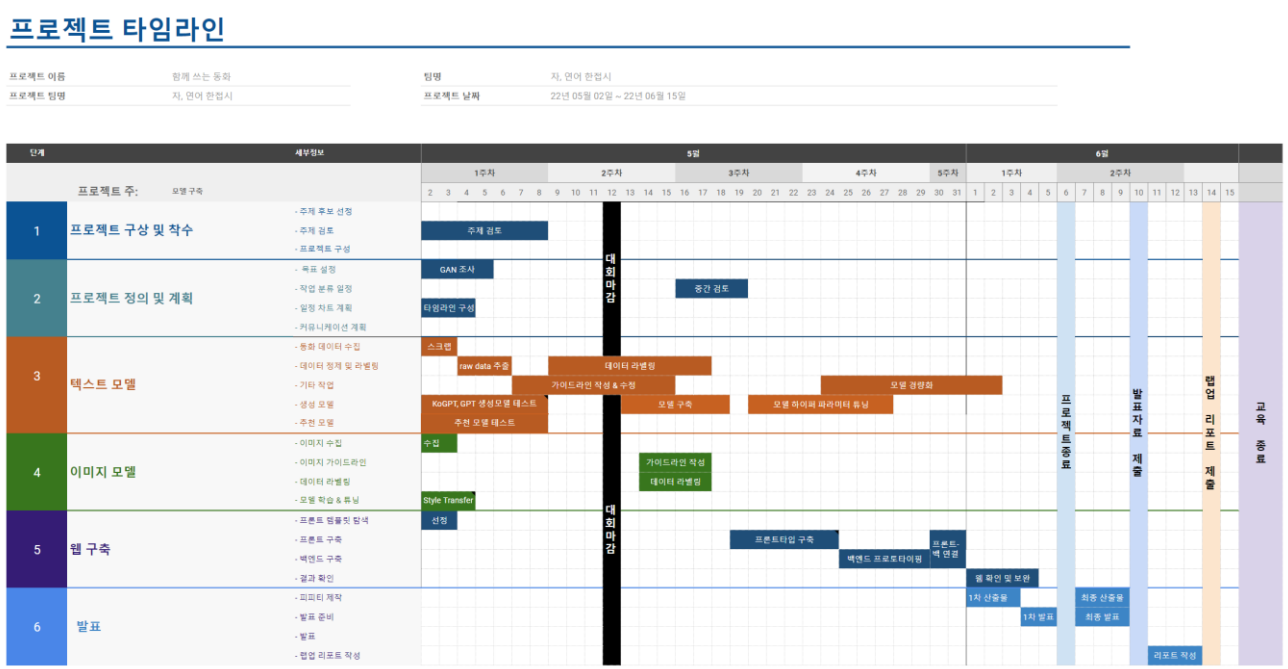


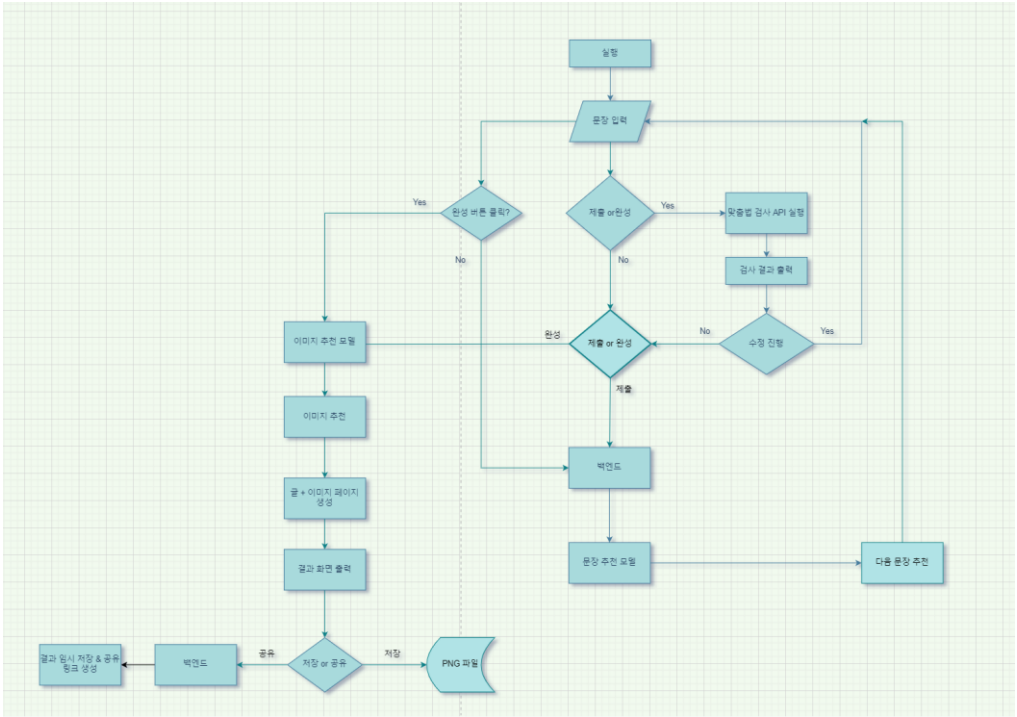


- Inference 이슈 & 코드 에러는 Github issue 사용
- 기능 및 모델 실험별 branch 생성 후 pull request

프로젝트 관리

- 프로젝트 일정 관리 ➔ excel & 구글 캘린더　　서비스 플로우 ➔ drawio





감사합니다

boostcamp ai tech