CONTENT IMPACT [StoryXAI] AI: Storyteller 프로젝트

# Story Keyword Maker API

스토리 생성 API & 콘텐츠 제작

## Contents

01

프로젝트 소개

- Story Keyword Maker API

02

DB 구축 및 API 개발

- DataBase 구축
- 알고리즘 구현 및 개발

03

시나리오 제작

- 강제결합법 이론
- 시나리오 성시경 완성

04

최종 결과물

- 영상 <성시경>
- 웹툰, 무빙툰 <성시경>

## 창작의 고통을 겪는 스토리텔러들

어떤 이야기 **소재**로 풀어쓰면 좋을까..

이 다음 **대사**로는 뭘 써야 하지? 마땅히 떠오르는 대사가 없는데…. 다 들어본 이야기인데 조금 신박한 아이디어 없을까?



작가, 크리에이터 혹은 <mark>스토리텔러</mark> 가 되고자 하는 일반인들에게

스토리 창작 어시스턴트 API

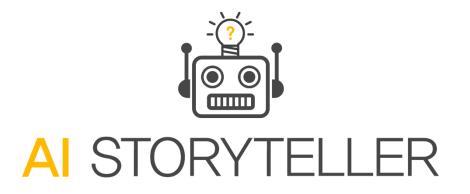
어떤 이야기 : Story Keyword Maker 를 제안한다!

다 들어본 이야기인데 조금 **신박한** 아이디어 없을까?

이 다음 대사로는 뭘 써야 하지? 마땅히 떠오르는 대사가 없는데….



## 스토리 창작 어시스턴트 API 구축

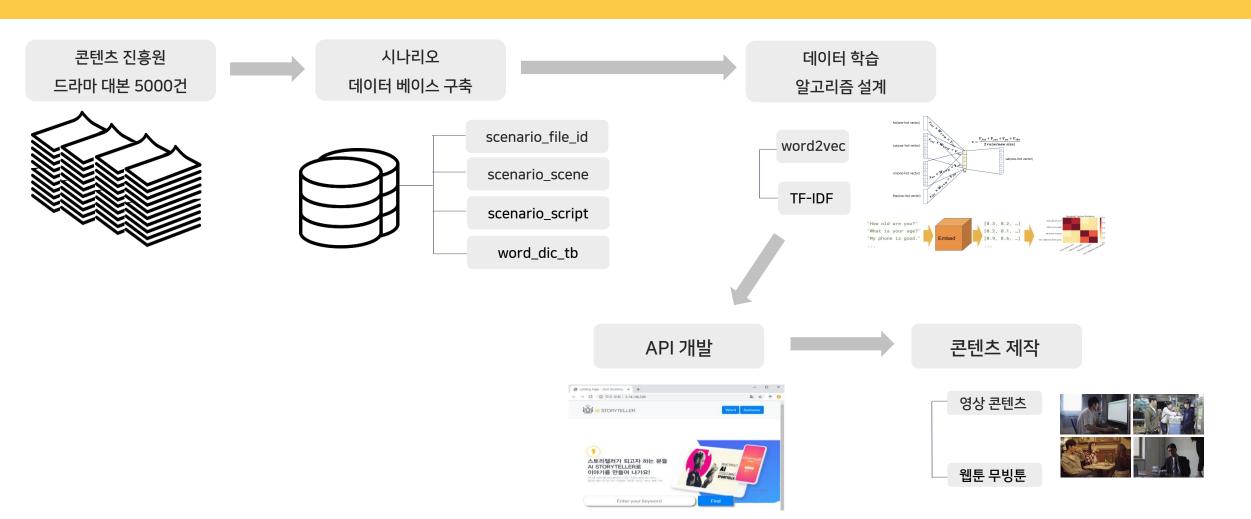


스토리 창작 어시스턴트 API 구축

스토리 관련 단어 입력했을 때 그 다음 키워드가 나오는 API 구현

스토리텔러가 되고자 하는 누구든 사용 가능 일반인 뿐만 아니라 작가 지망생도 새로운 시도로 서비스 활용 가능

## 프로젝트 진행 단계



## 시나리오 데이터 베이스 구축

드라마 데이터 50000건 중 28282건 데이터 베이스 구축

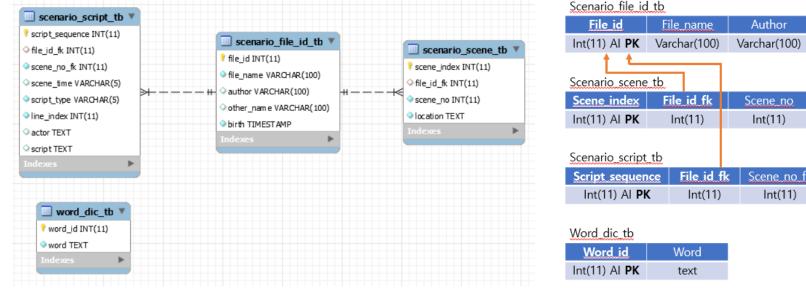
제공처	구분	건수	용량	
	문학	29건 (약 87만 어절)		
국립국어원	드라마 대본	19건 (약 69만 어절)	2GB	
	일상대화	294건 (약 266만 어절)		
	현대-문어 223건(약 771만 어절)			
한국방송작가협회	방송 대본	6만 건	5GB	
한국시나리오작가협회	시나리오	32건	3GB	
한국문예학술저작권협회	AJ	<u> 9만4천건</u>	30MB	

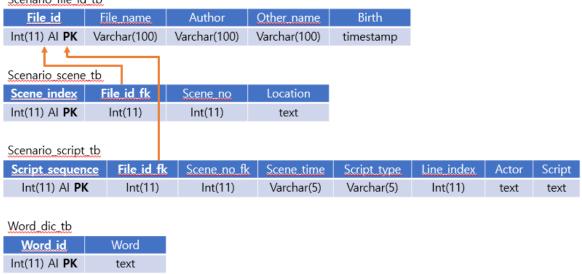
	file_id	file_name	author	other_name	birth
•	1	드라마시티 - 그는한강대교를가지않았다	가성진	M2007120100001	2019-10-25 03:20:57
	2	언제나두근두근 - 그녀, 볼수록 멋진 걸!!!	박영숙	M2007120100002	2019-10-25 03:20:58
	3	언제나두근두근 - 집으로 가는길	박영숙	M2007120100003	2019-10-25 03:20:59
	4	언제나두근두근 - 그 입술, 달콤한 첫사랑의	박영숙	M2007120100004	2019-10-25 03:21:01
	5	언제나두근두근 - 꿈★은 이루어질까?	박영숙	M2007120100005	2019-10-25 03:21:02
	6	언제나두근두근 - 첫사랑과 마지막사랑	박영숙	M2007120100006	2019-10-25 03:21:03
	7	언제나두근두근 - 내 연인의 약혼녀	박영숙	M2007120100007	2019-10-25 03:21:04
	8	언제나두근두근 - 연인들의 이별	박영숙	M2007120100008	2019-10-25 03:21:05
	9	언제나두근두근 - 너를 보내는 이유	가성진	M2007120100009	2019-10-25 03:21:06
	10	언제나두근두근 - 함께 하고 싶은 연인들	박영숙	M2007120100010	2019-10-25 03:21:07
	11	언제나두근두근 - 새로운 출발!	박영숙	M2007120100011	2019-10-25 03:21:08
	12	언제나두근두근 - 사랑은 가슴에 상처를 남	박영숙	M2007120100012	2019-10-25 03:21:09
	13	언제나두근두근 - 그대와 영원희	박영숙	M2007120100013	2019-10-25 03:21:10
	14	오픈드라마 남과여 - 조각난사랑	강남이	M2007120100014	2019-10-25 03:21:11
	15	금요일의천사 - 당신의 승리를 위해	강우진	M2007120100020	2019-10-25 03:21:12
	16	금요일의천사 - 동래성의守護神	강우진	M2007120100021	2019-10-25 03:21:12
	17	금요일의천사 - 無心川의절규	강우진	M2007120100022	2019-10-25 03:21:12
	18	금요일의천사 - 불멸의깃발	강우진	M2007120100023	2019-10-25 03:21:12
	19	금요일의천사 - 사랑스러운도둑	강우진	M2007120100024	2019-10-25 03:21:14
	20	금요일의천사 - 樂聖우륵	강우진	M2007120100025	2019-10-25 03:21:14
	21	금요일의천사 - 일그러진여름	강우진	M2007120100026	2019-10-25 03:21:14
	22	금요일의천사 - 큰별새벽에지다	강우진	M2007120100027	2019-10-25 03:21:14
	23	금요일의천사 - 표두장	강우진	M2007120100028	2019-10-25 03:21:14
	24	SBS오픈드라마남과여 - 해피버스데이1	김성희	M2007120100037	2019-10-25 03:21:14
	25	SBS오픈드라마 남과여 - 해피버스데이1부	강윤경	M2007120100037	2019-10-25 03:21:15

scenario\_file\_id 파일의 ID, 시나리오 이름, 작가, 콘진원에서 제공한 이름, 파일 생성 시간이 기록된 테이블

## 시나리오 데이터 베이스 구축

파일 복호화 과정 : 가상 환경 생성한 후 방송 대본 파일 가져옴





DataBase ER 다이어그램

## 시나리오 데이터 베이스 구축

각 씬, 대사, 어휘에 해당하는 데이터 테이블 생성

	scene_index	file_id_fk	scene_no	location
•	1	1	1	S# 1 촬영장
	2	1	2	S# 2 동 일각
	3	1	3	S# 3 촬영장 앞
	4	1	4	S#4패밀리 식당
	5	1	5	S# 5 한강대교위
	6	1	6	S# 6 퀵서비스 사무실(아침)
	7	1	7	S# 7 거리
	8	1	8	S# 8 골목길
	9	1	9	S# 9 양옥집 앞
	10	1	10	S# 10 골목길
	11	1	11	S# 11 진석의 자취방/부엌(
	12	1	12	S# 12 진석의 자취방
	13	1	13	S# 13 시장내의 먹자 골목(
	14	1	14	S# 14 음식점
	15	1	15	S# 15 거리(밤)
	16	1	16	S# 16 진석의 방
	17	1	17	S# 17 진석의 방(아침)
	18	1	18	S# 18 양옥집 앞
	19	1	19	S# 19 사무실
	20	1	20	S# 20 상가
	21	1	21	S# 21 스키용품점앞
	22	1	22	S# 22 상점안
	23	1	23	S# 23 진아의 회사 앞(저녁)
	24	1	24	S# 24 진아의 집 앞 (늦은밤)
	25	1	25	S# 25 거리(늦은밤)

	script_s	file_id_fk	scene_n	scene_time	script_type	line_index	actor	script
•	1	1	1	null	대사	1	감독(E)	엔지-
	2	1	1	null	지문	1		사람들, 동작 풀어지고 왜그런가 둘러보고, 한수도 이유
	3	1	2	null	대사	2	조감독	(작은 소리로)강선배, 미안해요. 감독이 빼라는데
	4	1	2	null	대사	3	한수	(기분 나쁘지만)왜에?
	5	1	2	null	대사	4	조감독	저기 그냥 똘마니 둘로 간대요.
	6	1	2	null	대사	5	한수	그래?
	7	1	2	null	대사	6	조감독	나 들어가우
	8	1	2	null	대사	7	한수	저기 다음에 필요하면 불러라
	9	1	2	null	대사	8	조감독	(들어가며)알았어요
	10	1	2	null	대사	9	한수	(뒤에다)부탁한다 (하면서 미련 남는 듯 촬영장 시선두
	11	1	2	null	지문	2		한수, 태호와 시선 부딪히고 태호, 팔짱낀채 냉정히 고개
	12	1	4	null	대사	10	진아(E)	하나아빠- 당신이 약속을 지키리라 기대도 하지 않았어
	13	1	5	null	대사	11	한수	(감정넣어 읽는다)여러분 감사합니다. 저에게 남우조연
	14	1	5	null	대사	12	한수(E)	누가 인생에서 주연이 아닌 조연으로 살다가길 바라는 사
	15	1	6	낮	대사	13	진석	이번달만 벌써 세껀이 넘네요?
	16	1	6	낮	대사	14	김사장	나참 아니, 본인한테 딱딱 전달하면 이런일이 일어나나
	17	1	6	낮	지문	3		한수, 살그머니 들어오고 있다.
	18	1	6	낮	대사	15	김사장	어이~ 뭐야?
	19	1	6	낮	대사	16	한수	(그제서야) 안녕하세요.
	20	1	6	낮	대사	17	김사장	이봐, 여기가 동네 놀이턴 풀 알아? 당신 나오고 싶으면
	21	1	6	낮	대사	18	한수	죄송합니다
	22	1	6	낮	대사	19	김사장	김실장이 소개해서 내가 잠자코 있을려고 했는데 당신
	23	1	6	낮	대사	20	진석	사장님, 제가 잘 얘기 하겠습니다.(한수 끌고 가는)
	24	1	6	낮	대사	21	김사장	(수그러들며 뒤에 대고)에이뭐 하나래도 똑부러지게 해
	25	1	6	낮	대사	22	진석	(김사장 안보이자)오늘도 갈거냐?

	word_id	word
•	1	
	2	엔지
	4	사람들
	5	동작
	6	풀어지고
	7	왜그런가
	8	둘러보고
	9	한수도
	10	이유몰라
	11	앞
	12	바라보면
	13	스텝들
	14	감독
	15	있고
	16	ユ
	17	옆으로
	18	태호
	19	제작자
	20	팔짱끼고
	21	보고있다
	1	

◆ 학습 데이터 생성 및 정제

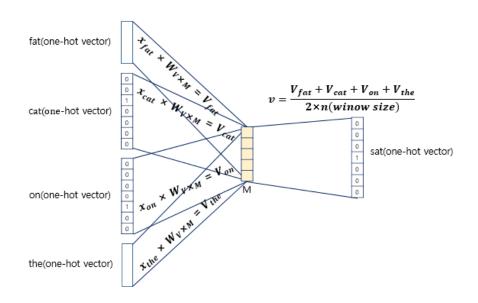
	script s	file_id_fk	scene n	scene time	script_type	line index	actor	script
•	1	1	1	null	대사	1	감독(E)	엔지-
	2	1	1	null	지문	1		사람들, 동작 풀어지고 왜그런가 둘러보고, 한수도 이유
	3	1	2	null	대사	2	조감독	· (작은 소리로)강선배, 미안해요. 감독이 빼라는데
	4	1	2	null	대사	3	한수	(기분 나쁘지만)왜에?
	5	1	2	null	대사	4	조감독	저기 그냥 똘마니 둘로 간대요.
	6	1	2	null	대사	5	한수	그래?
	7	1	2	null	대사	6	조감독	나 들어가우
	8	1	2	null	대사	7	한수	저기 다음에 필요하면 불러라
	9	1	2	null	대사	8	조감독	(들어가며)알았어요
	10	1	2	null	대사	9	한수	(뒤에다)부탁한다 (하면서 미련 남는 듯 촬영장 시선두
	11	1	2	null	지문	2		한수, 태호와 시선 부딪히고 태호, 팔짱낀채 냉정히 고개
	12	1	4	null	대사	10	진아(E)	하나아빠- 당신이 약속을 지키리라 기대도 하지 않았어
	13	1	5	null	대사	11	한수	(감정봉어 읽는다)여러분 감사합니다. 저에게 남우조연
	14	1	5	null	대사	12	한수(E)	누가 인생에서 주연이 아닌 조연으로 살다가길 바라는 사
	15	1	6	낮	대사	13	진석	이번달만 벌써 세껀이 넘네요?
	16	1	6	낮	대사	14	김사장	나참 아니, 본인한테 딱딱 전달하면 이런일이 일어나나
	17	1	6	낮	지문	3		한수, 살그머니 들어오고 있다.
	18	1	6	낮	대사	15	김사장	어이~ 뭐야?
	19	1	6	낮	대사	16	한수	(그제서야) 안녕하세요.
	20	1	6	낮	대사	17	김사장	이봐, 여기가 동네 놀이턴 줄 알아? 당신 나오고 싶으면
	21	1	6	낮	대사	18	한수	죄송합니다
	22	1	6	낮	대사	19	김사장	김실장이 소개해서 내가 잠자코 있을려고 했는데 당신
	23	1	6	낮	대사	20	진석	사장님, 제가 잘 얘기 하겠습니다.(한수 끌고 가는)
	24	1	6	낮	대사	21	김사장	(수그러들며 뒤에 대고)에이뭐 하나래도 똑부러지게 해
	25	1	6	낮	대사	22	진석	(김사장 안보이자)오늘도 갈거냐?

scenario\_script 대사 기준 데이터 테이블

script_sequence	script_type	script	clean_script	tokenized_script
1	지문	별이 가득한 광대한 우주공간 보이고 우주 선들 빠르게 날아가는 모습 보이다 공간 안 에	별이 가득한 광대한 우주공간 보이고 우주선 둘 빠르게 날아가는 모습 보이다 공간 안에	['별', '이', '가득한', '광대한', '우주', '공간', '보이', '고'
2	대사	(영어로 off) 준, 열 제어판 손상은 어느 정도 인가?	준 열 제어판 손상은 어느 정도인가	['준', '열', '제어판', '손상', '은', '어느', '정도', '인가']
3	대사	(영어) 조사중이다.	조사중이다	['조사', '중', '이다']
4	태사	(영어로) 육안으로 보이는 외상이 있는가?	육안으로 보이는 외상이 있는가	['육안', '으로', '보이는', '외상', '이', '있는가']
5	태사	(영어로) 없다. 그런데 아까 우리와 부딪친 게 뭐였나?	없다 그런데 아까 우리와 부딪친 게 뭐였나	['없다', '그런데', '아까', '우리', '와', '부딪친', '게', '뭐'
6	대사	(영어로 off) 인공위성 일부분. 2006년에 한 국에서 쓴 아리랑위성 잔해로 파악됐다.	인공위성 일부분 년에 한국에서 쏜 아리랑위 성 잔해로 파악됐다	['인공위성', '일부분', '년', '에', '한국', '에서', '쓴', '아리
7	대사	(영어로) 2006년? 골동품이네. (웃으며) 2006년이면 내가 태어난 핸데.	년 골동품이네 년이면 내가 태어난 핸데	['년', '골동품', '이네', '년', '이', '면', '내', '가', '태
8	대사	(영어로 off) 그래? (웃는) 이상.	그래 이상	['그래', '이상']
9	대사	Roger. (교신 끝내고는) 2006년? 하(고개 를 돌려 어딘가를 바라보는)	년 하	['년', '하']
10	지문	카메라 팬하며 우주공간에서 지구로, 지구 에서 아시아 대륙, 아시아에서 한국으로 빠 른	카메라 팬하며 우주공간에서 지구로 지구에 서 아시아 대륙 아시아에서 한국으로 빠른 디	['카메라', '팬', '하며', '우주', '공간', '에서', '지구', '로'
11	대사	(NARR) 지금으로부터 28년전인 2006년 내가 태어난 해다대한민국	지금으로부터 년전인 년내가 태어난 해다대 한민국	['지금', '으로부터', '년', '전인', '년내', '가', '태어난', '해
12	지문	빠른 디졸브로 내레이션에 관련된 여러 자료화면 나가고. (노무현 대통령활우석 박	빠른 디졸브로 내레이션에 관련된 여러 자료 화면 나가고	['빠른', '디졸', '브로', '내레이 선', '에', '관련', '된', '여러
13	대사	(NARR) 대한민국의 2006년은 노무현대통 령 5년임기의 후반무렵으로 격동의 한해	대한민국의 년은 노무현대통령 년임기의 후 반무렵으로 격동의 한해였다 줄기세포조작 사	['대한민국', '의', '년', '은', '노 무현', '대통령', '년', '임기
14	지문	카메라, 흑석동을 부감으로 비춘다.	카메라 흑석동을 부감으로 비춘다	['카메라', '흑석동', '울', '부감', '으로', '비춘다']
15	대사	(NARR) 그 격동과 좌절의 2006년 늦여름 나는 서울 흑석동에서 태어났다.	그 격동과 좌절의 년 늦여름 나는 서울 흑석동 에서 태어났다	['그', '격동', '과', '좌절', '의', '년', '늦', '여름', '나
16	지문	화면 블랙에서 밝아지면 아기 몽(준이)의 시 선컷으로 가족들 얼굴 보인다. 문희와 해	화면 블랙에서 밝아지면 아기 몽의 시선컷으로 가족들 얼굴 보인다 문희와 해미 얼굴이	['화면', '블랙', '에서', '밝아지 면', '아기', '몽', '의', '시선

전처리 과정: 토큰화 및 불용어 제거

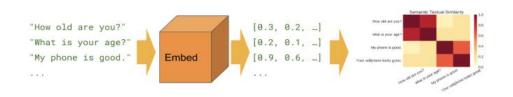
API 알고리즘 설계



#### Word2vec 임베딩 모델

'비슷한 위치에서 등장하는 단어들은 비슷한 의미를 가진다' 주어진 주변 단어를 통해 중심 단어를 예측하는 과정을 거침 이 학습 과정 통해 단어는 벡터 사상에 맵핑되고, 단어가 벡터로 구현됨으로써, 각 단어 간의 유사도 파악이 가능해짐

#### ◆ API 알고리즘 설계



$$W_{x,y} = tf_{x,y} \times log(\frac{N}{df_x})$$

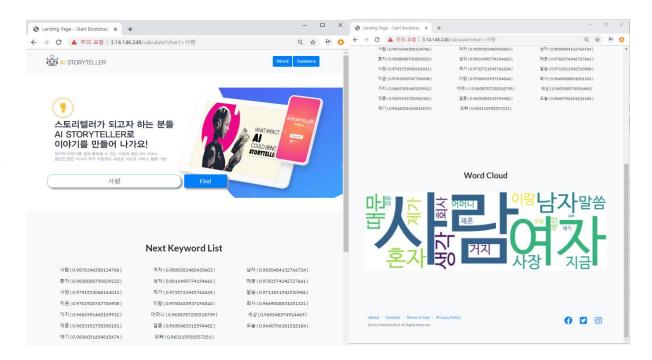


 $tf_{x,y}$  = frequency of x in y  $df_x$  = number of documents containing x N = total number of documents

#### TF-IDF 모델

문서 내 상대적으로 중요한 단어에 가중치를 부여하는 방식 Tf 단어 빈도, idf 문서 빈도의 역수 특정 문서 내에서 단어 빈도가 높으며 동시에 전체 문서에서 적게 포함되면 이 값은 높아지는 개념 문서 유사도에서 자주 사용되는 기법

Word2vec 임베딩 기법 , TF-IDF 코드 구현 및 데모 페이지 생성



모델 코드 구현

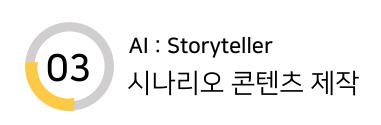
Story Keyword Maker API 웹 페이지 생성

## API를 활용해 시나리오 생성

강제 결합법 (forced relationship)

전문 스토리 창작자이라도 창작 과정에서 단어 선택, 사건 전개 방법 선택에 어려움을 겪음

강제 결합법(forced relationship)은 1958년에 영국의 작가인 찰스 화이팅(Charles S. Whiting)이 개발한 사고 기법으로, 전혀 관계가 없어 보이는 두 가지 이상의 아이디어나 사물에 대한 정보를 제시하면 억지로 연결시켜서 새로운 아이디어를 산출할 수 있도록 돕게 됨



## API + 강제 결합법을 활용해 이야기 생성

기존 시나리오 작법과는 다른 방식으로 접근 - 단어 기준으로 이야기 생성

주인공의 목표: 입사

장애물 : 블라인드 평가



입사	블라인드	부정	조작	양기탁	침착
하하하	비리	후후	미야	독립	대리인
결정	사항	실력	체포	사항	통과
후후	응급	하하하	가격	응급	사항
무례	에선	경력	모나리자	호흡기	이등
퇴원	내부	마찬가지	설치	비리	물산
와이프	양기탁	정직	왼쪽	가격	응급
사고	대리인	삼천	응급	에선	비리
부정	조작	해주시	양기탁	물산	현상
5년	물산	대한민국	오른쪽	통과	독립
운명	통과	부인	대상	침착	상대로

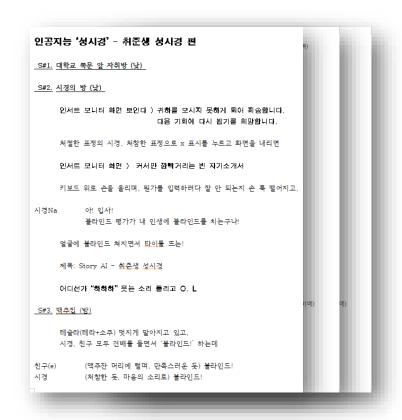
## API + 강제 결합법을 활용해 이야기 생성

인공지·	능 '성시 <sup>:</sup>	경' - 취준생				
입사	블라인드					
1 . 하하하	1 . 비리					
2 . 결정	2 . 사항	시경	입사	할 수 있 <del>을</del> 까?		
3 . 후후	3 . 응급	친구	<mark>1 . 하하하</mark>	니가?		
4 . 무례	4 . 에선	시경	<mark>블라인드</mark>	란다!		
5 . 퇴원	5 . 내부	친구	<mark>2 . 사항</mark>	2 . 결정	을 해야 한다.	
6 . 와이프	6 . 양기탁	시경	1 . 비리	밖에는 답이 없다.		
7 . 사고	7 . 대리인	친구	3 . 응급	이다.		
8 . 부정	8 . 조작	시경	후후			
9 . 5년	9 . 물산	친구	5 . 내부	4 . 에선		
10 . 운명	10 . 통과	시경	<mark>6 . 와이프</mark>	5 . 퇴원		
		친구	<mark>6 . 양기탁</mark>	이라고 했지?		
		시경	<mark>7 . 사고</mark>	다.		
		친구	7 . 대리인	8 . 부정	8 . 조작	을 해서라도
		시경				
		친구	9 . 5년	10 . 운명	이라고 생각해?	
		시경	<mark>9 . 물산</mark>	은	10 . 통과	래!
			7 . 대리인	8 . 부정	8 . 조작	안된다!

인공지능 '성시경' - 취준생 편							
부정	조작						
후후	□ 0 <b></b>						
2 . 실력	2 . 체포	시경		부정	할 수 있을까?		
3 . 하하하	3 . 가격	친구		후후	니가?		
4 . 경력	4 . 모나리자	시경		조작	란다!		
5 . 마찬가지	5 . 설치	친구		2 . 체포	2 . 실력	을 해야 한다.	
6 . 정직	6 . 왼쪽	시경		미야	밖에는 답이 없다	<b>ŀ</b> .	
7 . 삼천	7 . 응급	친구		3 . 가격	이다.		
8 . 해주시	8 . 양기탁	시경		후후			
9 . 대한민국	9 . 오른쪽	친구		5 . 설치	4 . 모나리자		
10 . 부인	10 . 대상	시경		6 . 정직	5 . 마찬가지		
		친구		6 . 왼쪽	이라고 했지?		
		시경		7 . 삼천	다.		
		친구		7 . 응급	8 . 해주시	8 . 양기탁	을 해서라도
		시경		??			
		친구		9 . 대한민	10 . 부인	이라고 생각해라	-
		시경		9 . 오른즉	은	10 . 대상	래!
				7 . 응급	8 . 해주시	8 . 양기탁	안된다!

## 인공지능 + 작가 + 연기자가 함께 만들어가는 스토리텔링

인공지능 "성시경" - 취준생 성시경 편

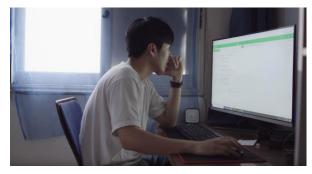


창작	Story Al	변문경 작가	연기자 창작
대사	60%	20%	20%
지문	10%	70%	20%

- 13 개의 씬으로 구성된 짧은 시나리오 완성
- 작가와 API 비중을 조절하며 다양한 이야기 생성 시도
- 연기자 또한 스토리를 읽고 해석하며 스토리 생성의 한 주체로서 역할을 함



#### Story Keyword Maker API - 영상 콘텐츠 제작









#### 블라인드 (6m10s)

각본: Story keyword Maker API, 변문경 작가

영상 제작 : 스튜디오 셰퍼드

연출 : 이민섭

제작 : 문진경

촬영 : 한동헌

AI: Storyteller 최종 결과물

Story Keyword Maker API - 웹툰, 무빙툰 콘텐츠 제작





치는구나!





취준생 고시경 (1화~3화)

각본: Story keyword Maker API, 변문경 작가

웹툰 제작, 연출: 이준복

그림 : 제이케이



## 스토리 창작 어시스턴트 API 여러분들의 창작활동에 도움이 되고 싶습니다.

#### Story Keyword Maker API



#### 웹드라마



웹툰, 무빙툰





# 감사합니다.