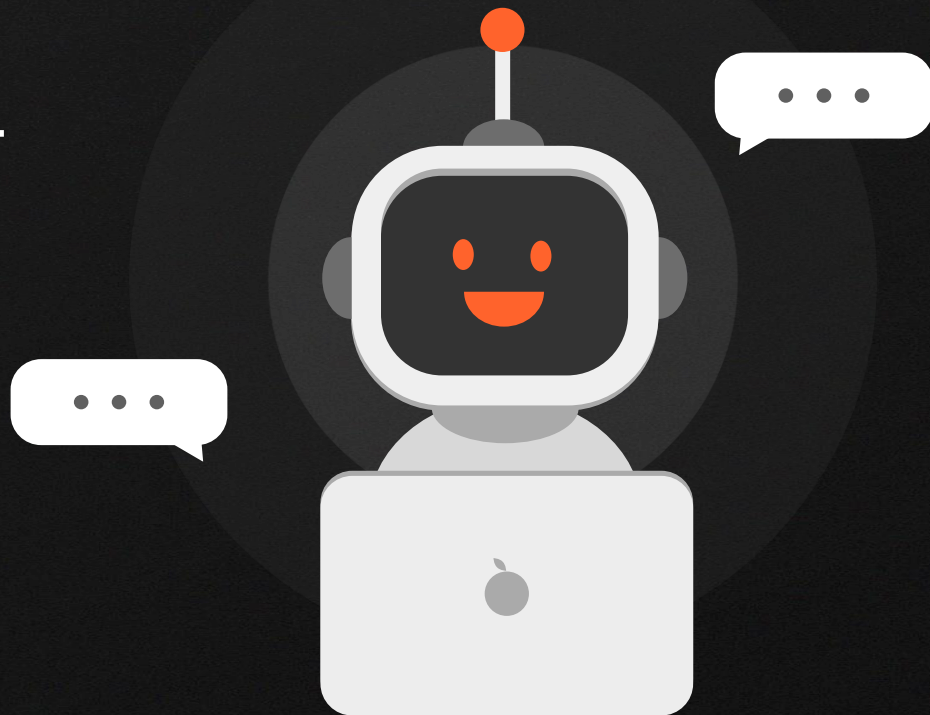


#자연어

#대화
시스템

자연어 처리 실습 : CHATBOT

김원우 강사



CONTENTS



CHAPTER 1

자연어 처리

01. 자연어 처리란?

02. 자연어 이해 및

자연어 생성의
정의

1) 자연어 이해 예시

2) 자연어 생성 예시

CHAPTER 2

대화 시스템

01. 대화 시스템이란?

02. 목적 지향 대화 시스템

1) 자연어

이해

2) 대화 관리

3) 자연어

생성

03. 오픈 도메인 대화 시스템

1) 검색 기반

방식

2) 생성 기반

방식

CHAPTER 3

실습

01. 목적 지향 대화 시스템

1) 자연어 이해 실습

2) 대화 관리 및 자연어 생성
실습

02. 오픈 도메인 대화 시스템

1) 생성 기반 방식
실습

03. 하이브리드 대화 시스템

1) 시스템 구성 및
실습

CHAPTER 1

자연어 처리

자연어 처리란?

$$\text{NLP}_{\textcircled{1}} = \text{NLU}_{\textcircled{2}} + \text{NLG}_{\textcircled{3}}$$

- ① NLP(자연어 처리) : Natural Language Processing
- ② NLU(자연어 이해) : Natural Language Understanding
- ③ NLG(자연어 생성) : Natural Language Generation

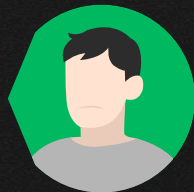
자연어 이해 및 자연어 생성의 정의



자연어 이해(NLU) 예시

문장

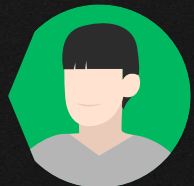
의미



영화 리뷰
시간이 남아도는 사람만 이 영화를
보세요.



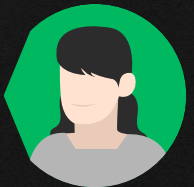
부정적 리뷰



유튜브 댓글
아니 팬 아닌데 방탄소년단 춤 왜캐
잘추냐



긍정적 댓글



상품 리뷰
옷이 핏하게 잘 떨어지네요~!



긍정적 리뷰

자연어 생성(NLG) 예시

의미

도움이 필요한 사람에게 위로 해주기



문장

많이 힘들겠구나, 내가 위로해 줄게



달고나 판매를 위한 마케팅 문구 생성



오징어 게임에 나온 그 달고나!



코로나 19 확진자 수를 알려주는 뉴스 제목 생성



오늘 코로나 19 확진자 수는 1230명 돌파, 역대 최고..



각 분야 활용 예시

과제	벤치마크	예제	
감정 분석	SST, IMDB, NSMC	입력	문장: 신만이 이 영화를 용서할 수 있다
		출력	긍정, 부정
유사도 예측	STS, MRPC, QQP, PAWS-X	입력	문장1 파이썬과 자바 중 뭐부터 배워야 하나요? 문장2 JAVA 나 Python 중 하나를 배워야 한다면, 뭐부터 시작해야 할까요?
		출력	유사, 비유사
자연어 추론	SNLI, MNLI, XNLI, aNLI	입력	전제: 회색 개가 숲에서 쓰러진 나무를 핥고 있다.
		출력	가설: 개가 밖에 있다. 참, 거짓, 모름

각 분야 활용 예시

과제	벤치마크	예제	
기계 독해	SQuAD, RACE, MS MARCO, KorQuAD	입력	지문 : 시카고 대학의 학자들은 경제학, 사회학, 법학 분석, 문학 비평, ... 이 대학은 미국 최대 대학 출판사인 시카고 대학 프레스의 본거지이기도 하다. 오는 2020년 완공될 버락 오바마 대통령 센터에는 오바마 대통령 도서관과 오바마 재단 사무소가 세워질 예정이다. 질문 : 버락 오바마 대통령 센터의 완공 예정연도는?
		출력	2020년
의도 분류	ATIS, SNIPS, AskUbuntu	입력	질문 : 현재 위치 주변에 있는 맛집 두 곳만 알려줘
		출력	장소검색 (식당, 가까운, 평점 좋은, 2)
언어적 용인 가능성	CoLA	입력	문장: 그는 한 번도 프랑스에 갔다.
		출력	가능, 불가능

NLP

NLU

의미론적
유사도 측정

감정
분석

기계
독해

목적 기반
대화

의미론적
구문 분석

관계
추출

상식
추론

생성형
문서 요약

생성형
질의 답

E2E
챗봇

NLG

자동
완성

스토
리 생성

생성형
언어 모델

캡션
생성

데이터 기반
문장 생성

표제어
추출

형태소
분석

구문
분석

기계
번역

CHAPTER 2

대화 시스템

대화시스 템

(Dialogue system)

1

목적 지향 대화 시스템

(Task-oriented Dialogue System)

특정한 목적, 목표를 위한
대화를 하는 시스템

2

오픈 도메인 대화 시스템

(Open domain Dialogue System)

특정한 목적, 목표 없이 자유로운
주제의 대화를 하는 시스템

대화 시스템 예시

목적 지향 대화 시스템

(Task-oriented Dialogue System)

SKT NUGU 인포콜

SKT NUGU 백신 케어콜

KT AICC

NAVER Ai Call

NAVER Happy Call

오픈 도메인 대화 시스템

(Open domain Dialogue System)

스캐터랩 이루다

심심이

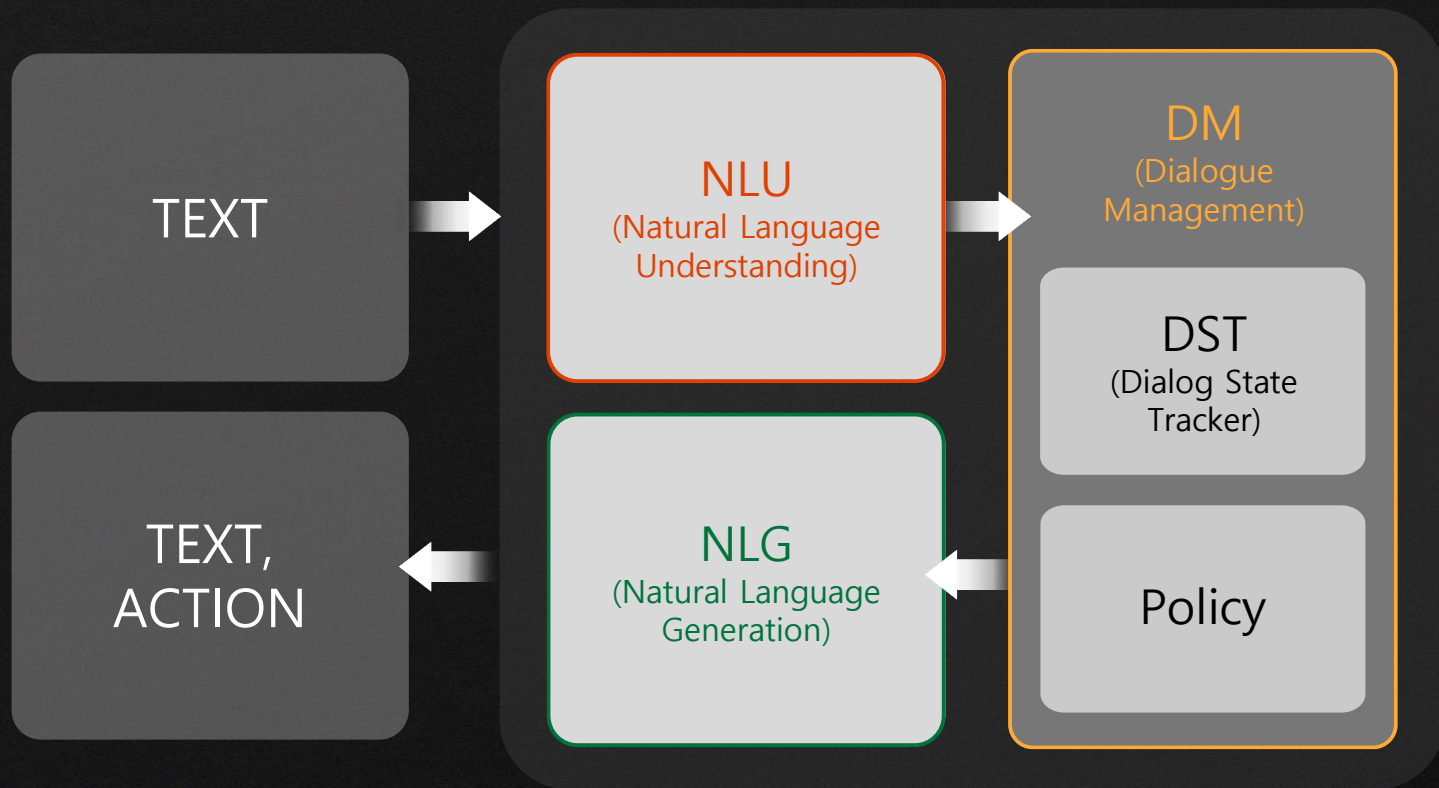
SKT NUGU Naver CLOVA 카카오 미니 삼성 빅스비 Apple siri KT GIGA
genie

1

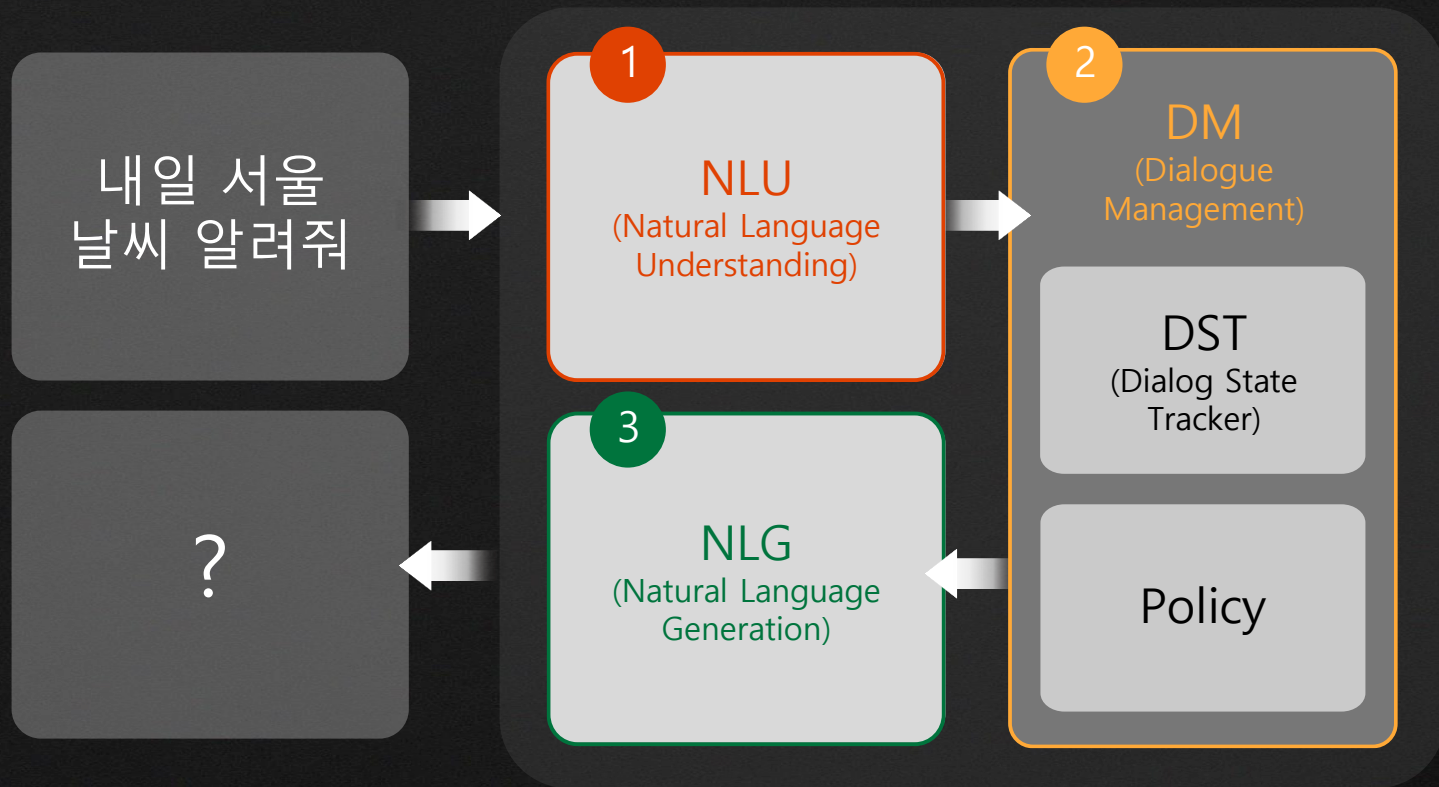
목적 지향 대화 시스템

(Task-oriented Dialogue System)

목적 지향 대화 시스템(Task-oriented Dialogue System)



목적 지향 대화 시스템(Task-oriented Dialogue System)



1.NLU(Natural Language Understanding)



The diagram consists of two main circular components side-by-side. The left component is orange and contains the text 'Intent' and '의도 파악'. The right component is green and contains the text 'Slot' and '정보 추출'. Both components are composed of three concentric circles, with the innermost circle being the most vibrant color and the outer rings being progressively darker shades of the same color.

Intent
의도 파악

Slot
정보 추출

1.NLU(Natural Language Understanding)

“ 내일 서울 날씨 알려줘 ”

Intent (의도)

Inform(Weather)

1.NLU(Natural Language Understanding)

“ 내일 서울 날씨 알려줘 ”

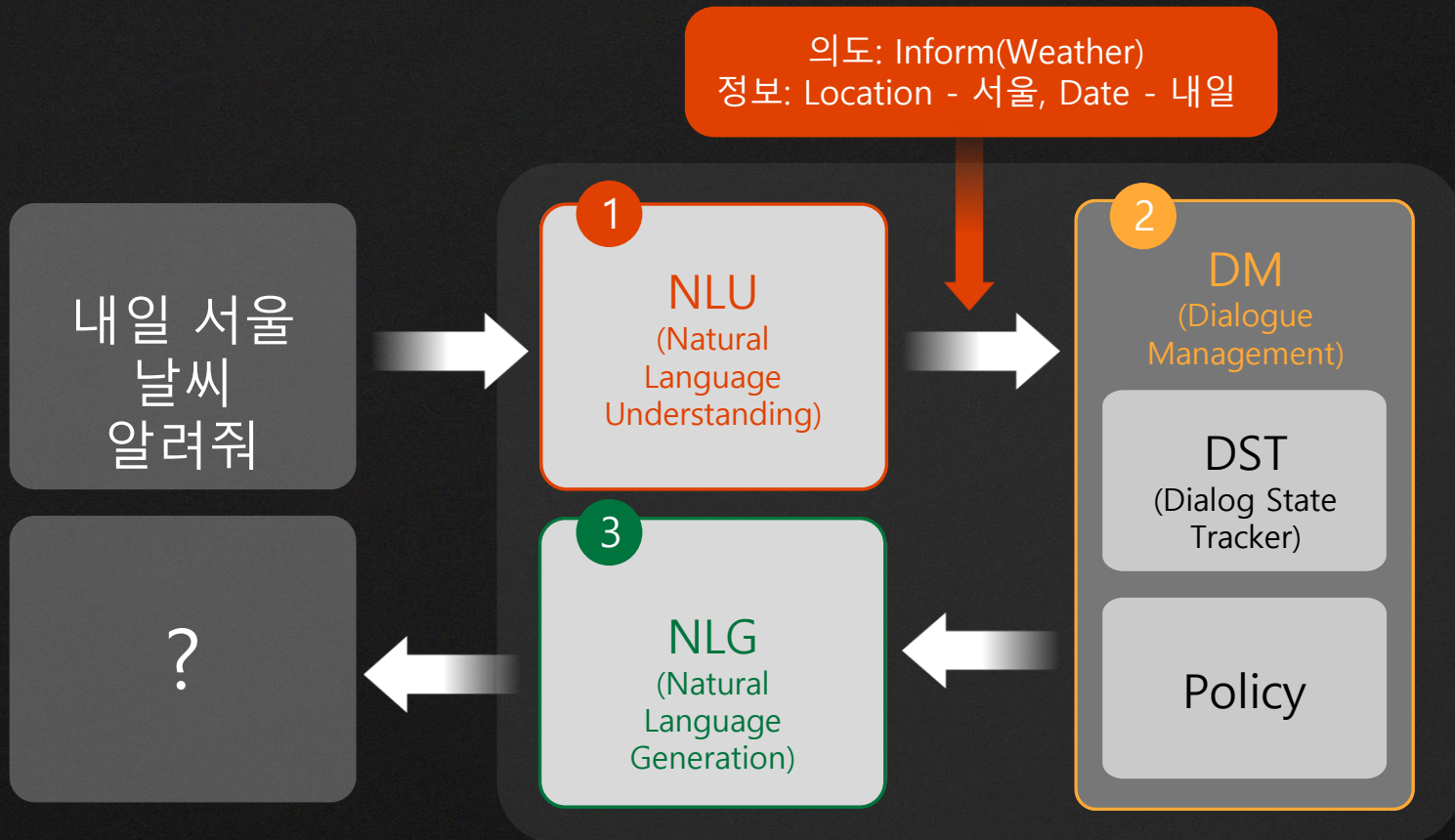
Intent (의도)

Inform(Weather)

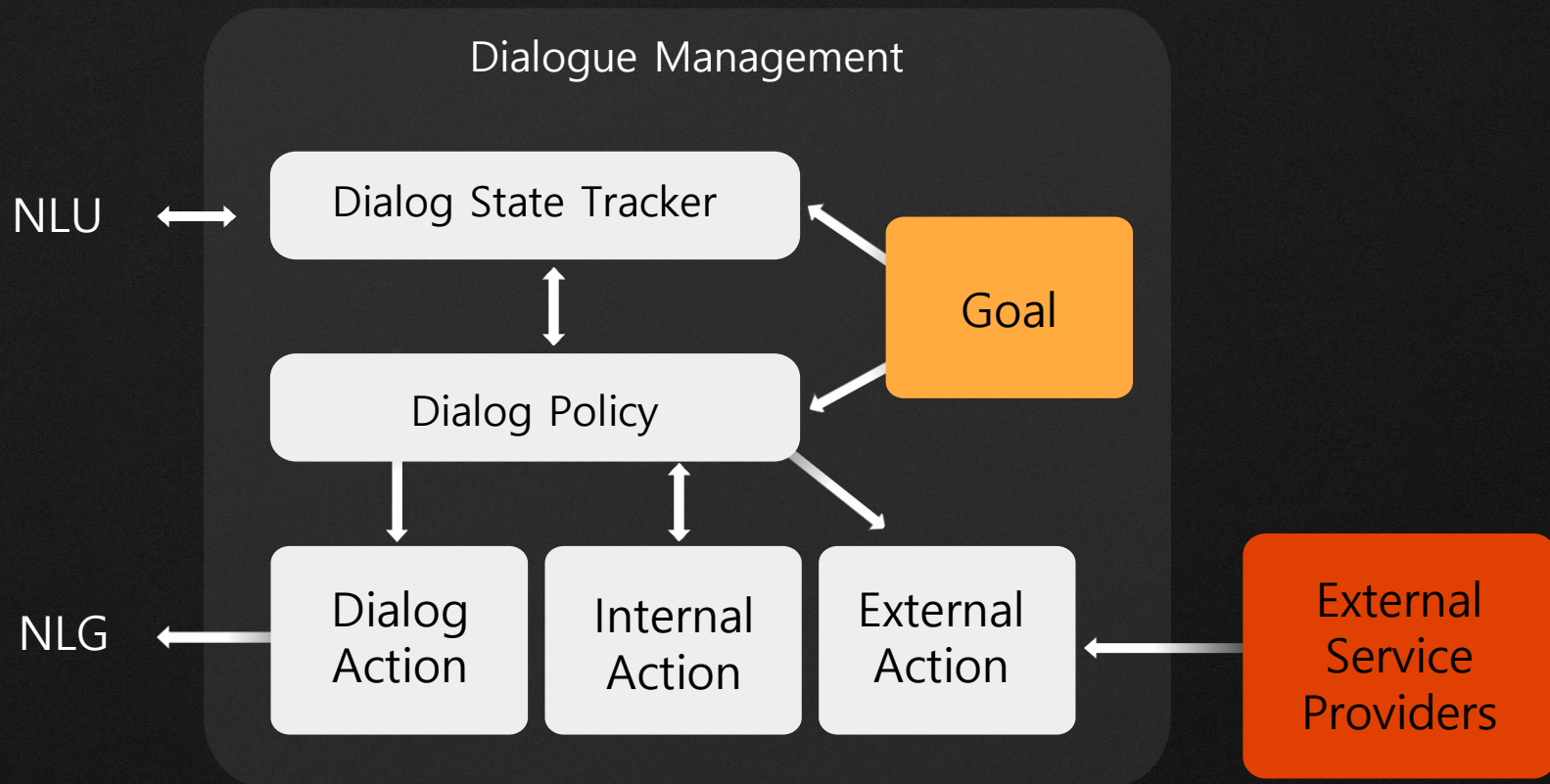
Slot(정보)

Location: 서울
Date: 내일

목적 지향 대화 시스템(Task-oriented Dialogue System)

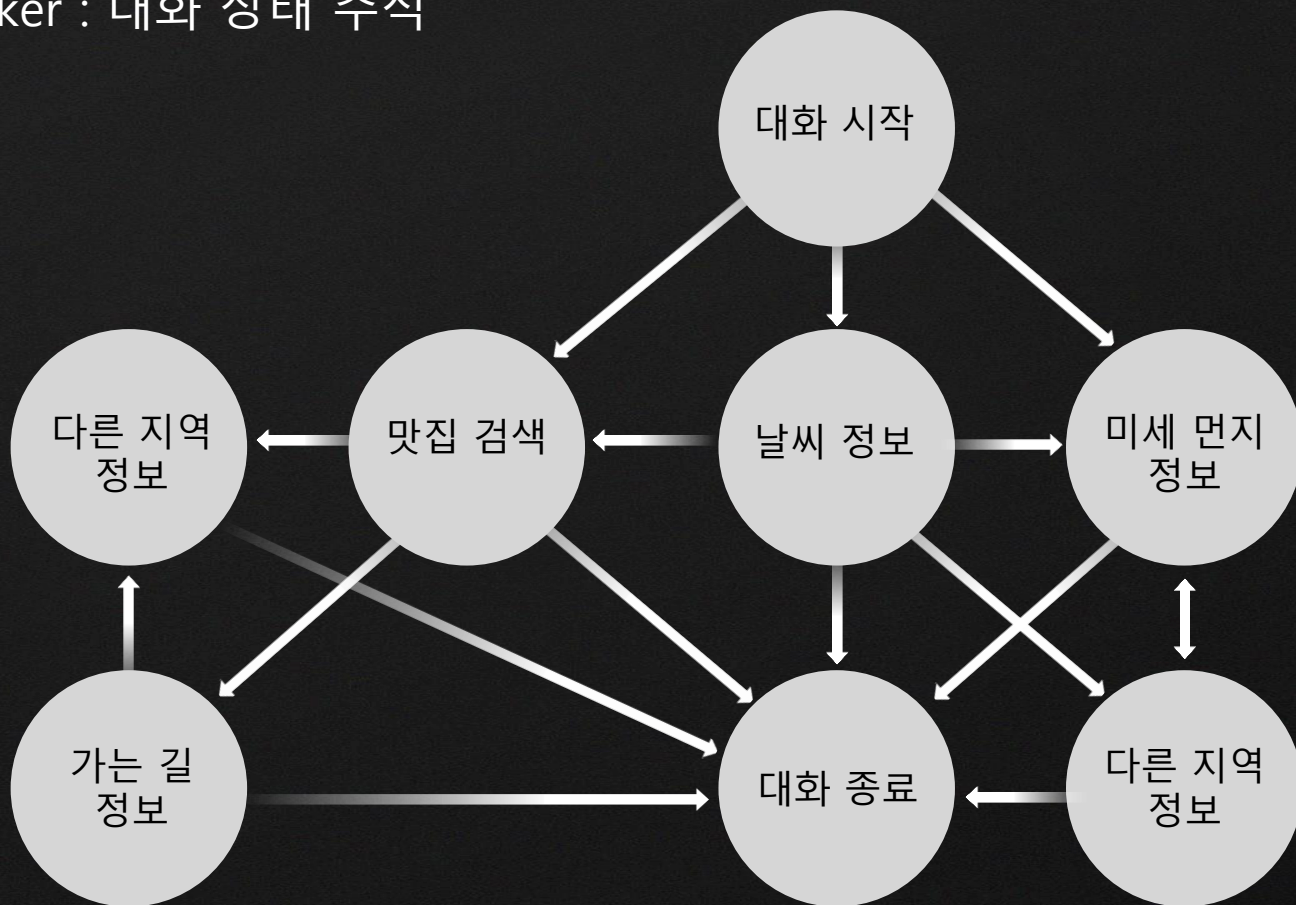


2. 대화 관리 (Dialog Management)



Dialog State Tracker : 대화 상태 추적

대화의 흐름

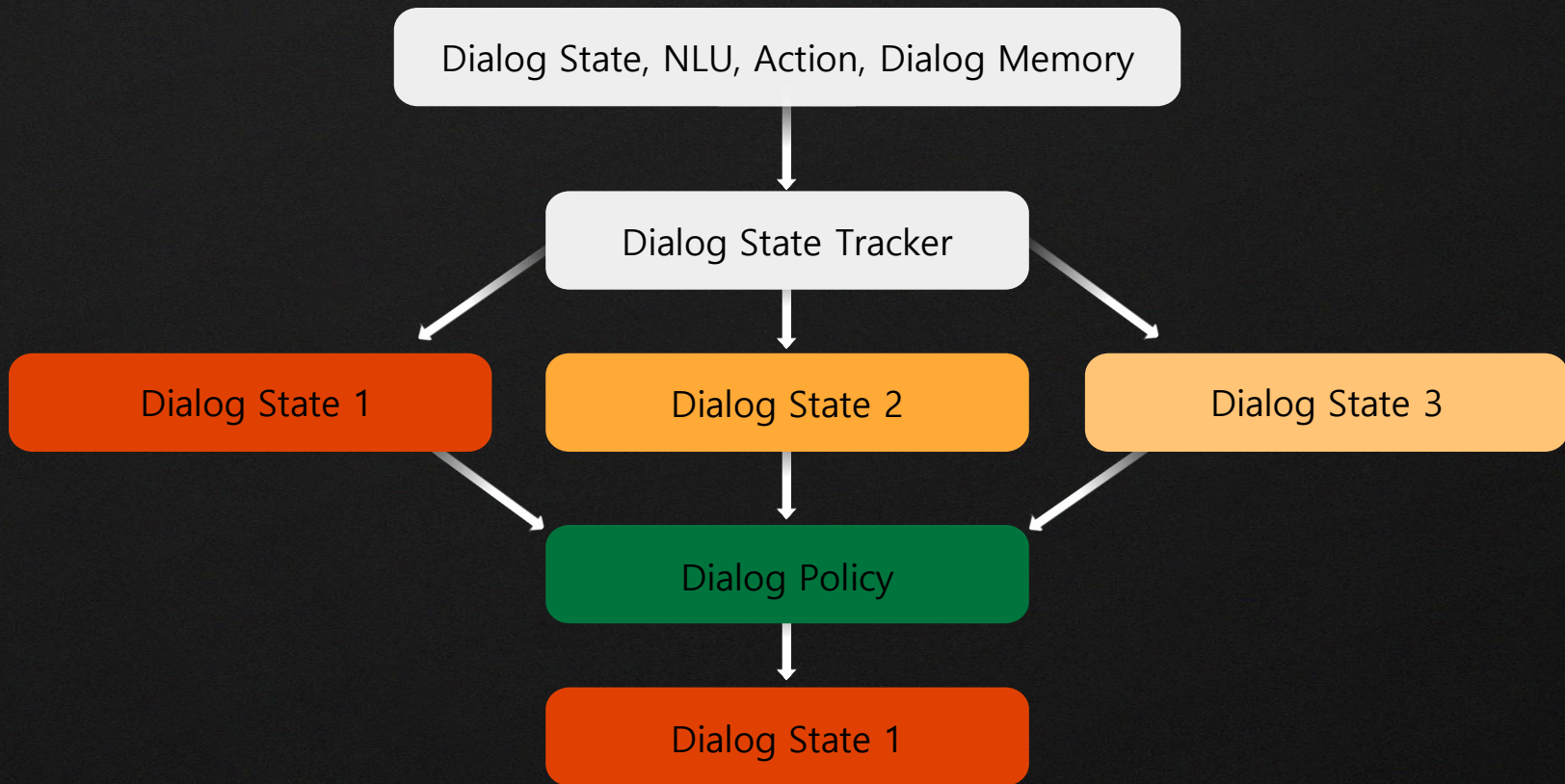


Dialog State Tracker : 대화 상태 추적

대화 이력 : 내일 날씨 알려줘 > 대전도 알려줘



Dialog Policy : 대화 정책



Dialog Policy : 대화 정책

사용자 발화 : 대전도 알려줘

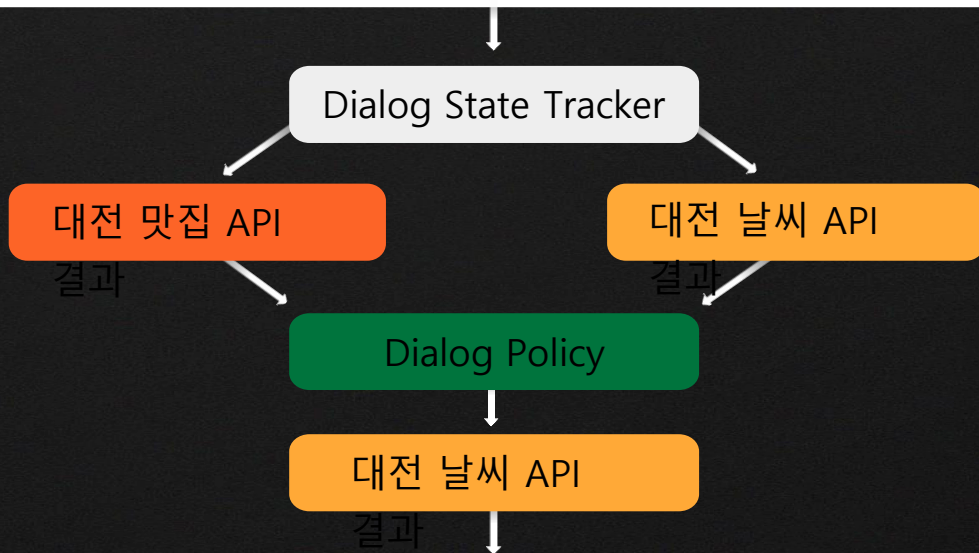
Dialog State : 날씨 정보

NLU : intent – Inform, slot – location(대전)

Action : 날씨 API, 미세먼지 API, 맛집 API

Dialog Memory : 날씨정보(location - 서울,

Date - 내일, weather_status -맑음)



" 내일 대전 날씨는 구름이
많습니다. "

3.자연어 생성 (Natural Language Generation)

NLG 템플
릿

@DATE @LOCATION 날씨는 @WEATHERSTATUS 입니다.

3.자연어 생성 (Natural Language Generation)

“내일 서울 날씨
알려줘”

NLU slot: Location - 서울, Date - 내일

DM External Action Result (날씨 API): 맑음

@DATE @LOCATION 날씨는 @WEATHERSTATUS 입니다.

3.자연어 생성 (Natural Language Generation)

“내일 서울 날씨
알려줘”

NLU slot: Location - 서울, Date - 내일

DM External Action Result (날씨 API): 맑음

내일

서울

날씨는

맑음

입니다.

2

오픈 도메인 대화 시스템

(Open Domain Dialogue System)

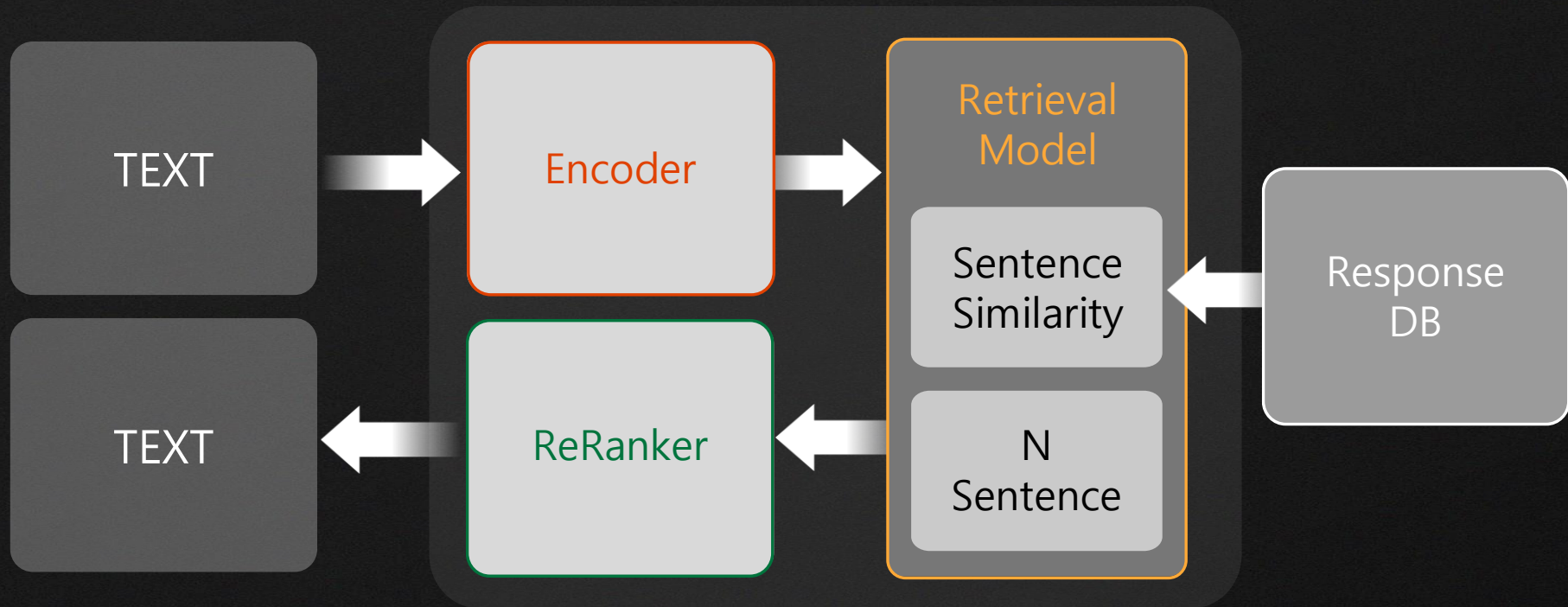
오픈 도메인 대화시스템

(Open Domain Dialogue System)

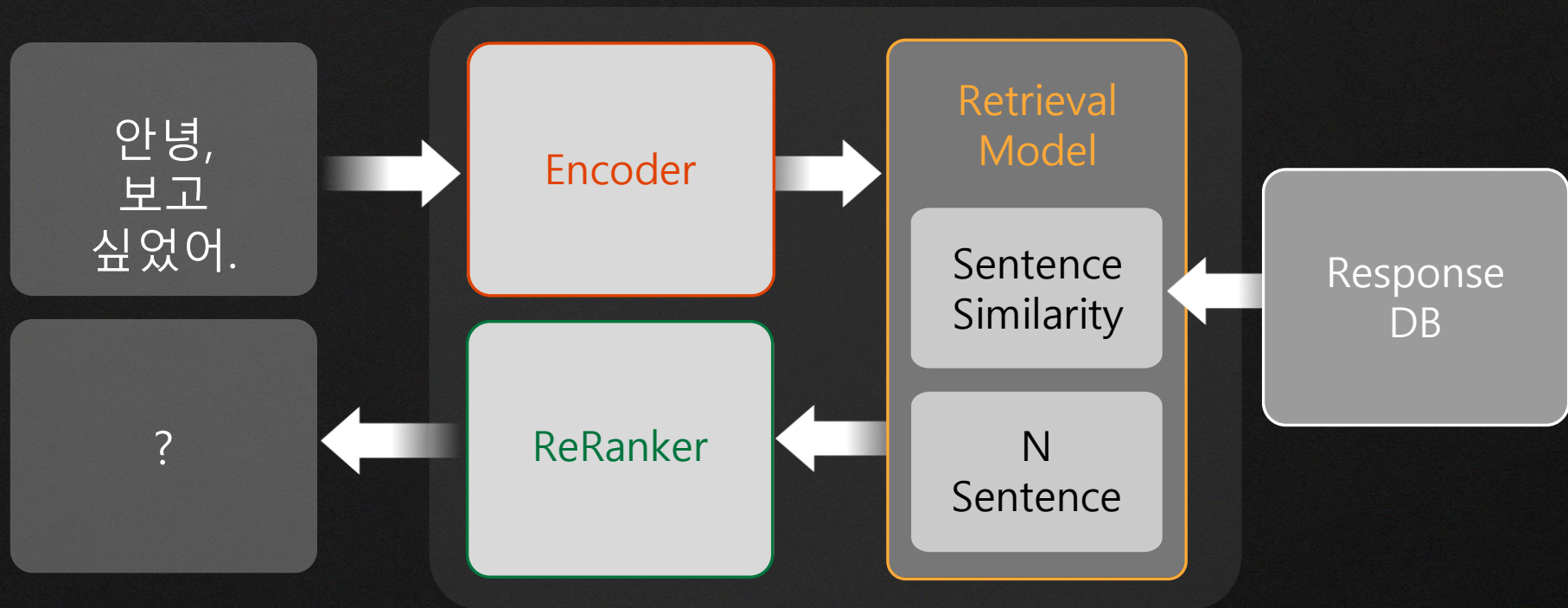
답변 검색기반 방식
(retrieval based)

답변 생성기반 방식
(generative based)

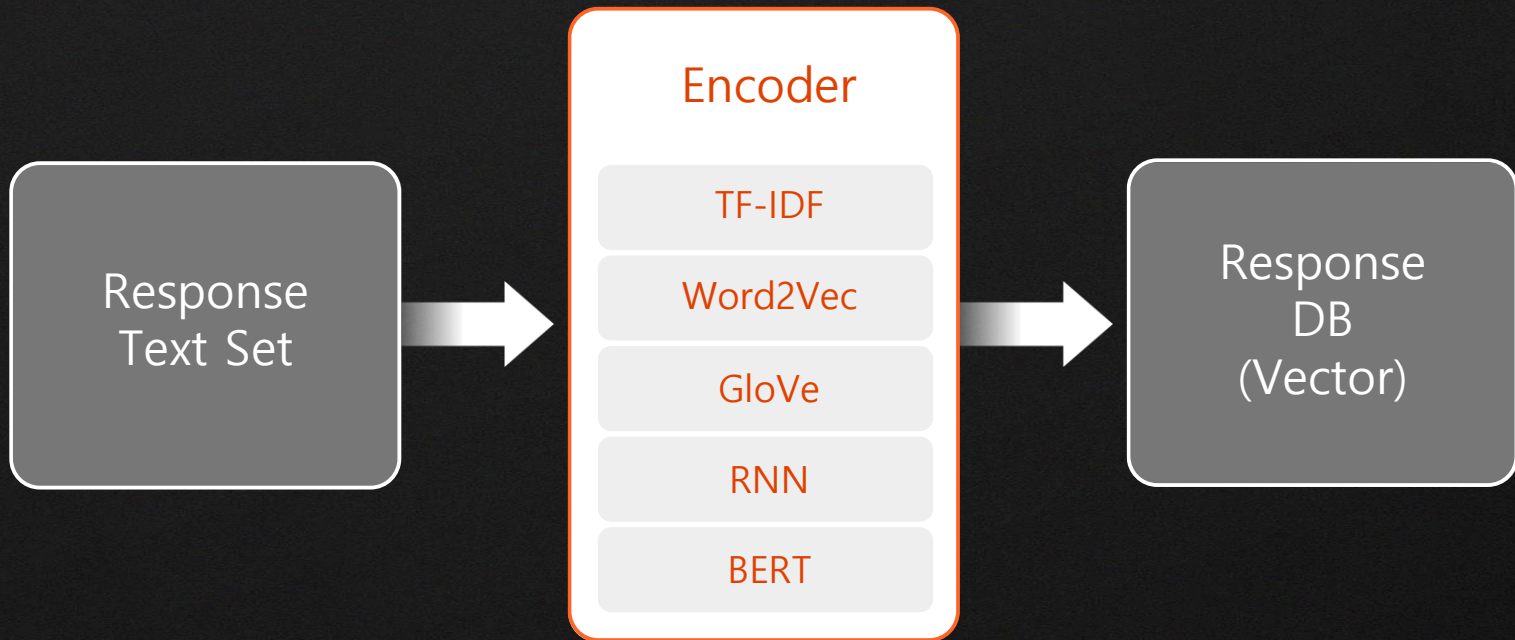
검색기반 방식



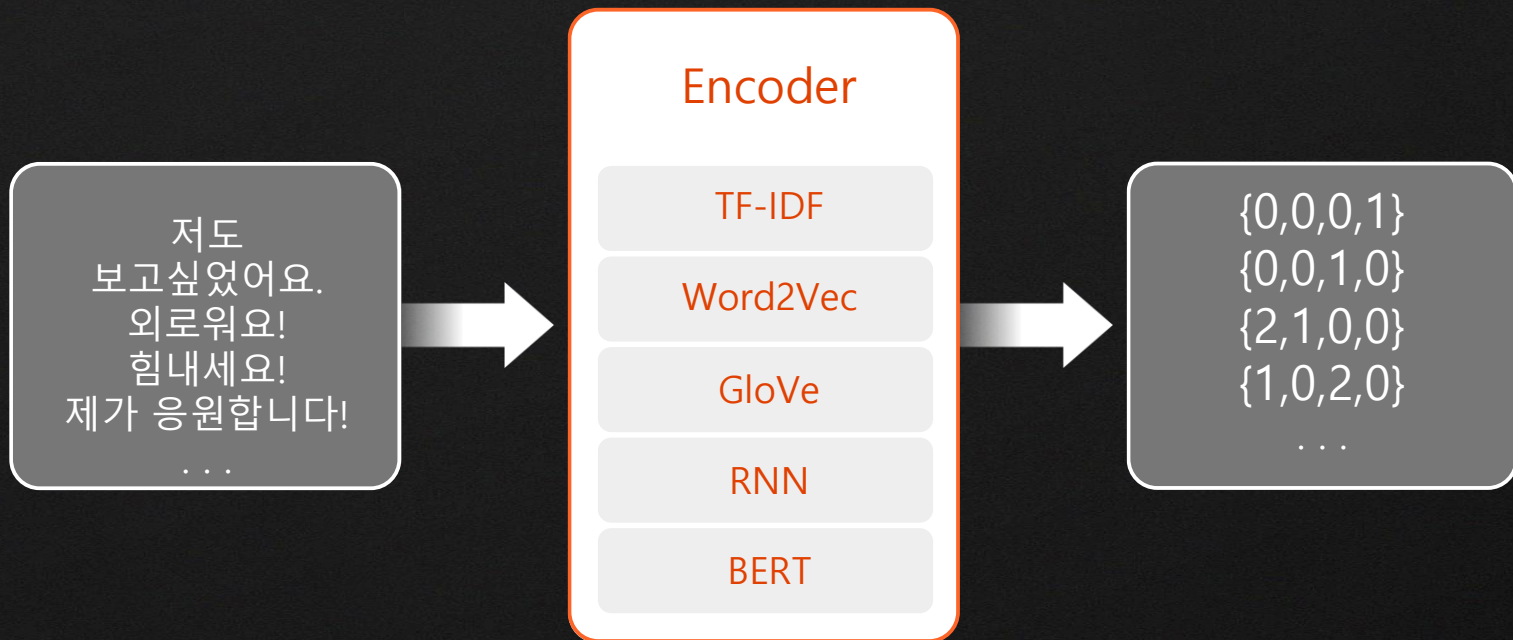
검색기반 방식



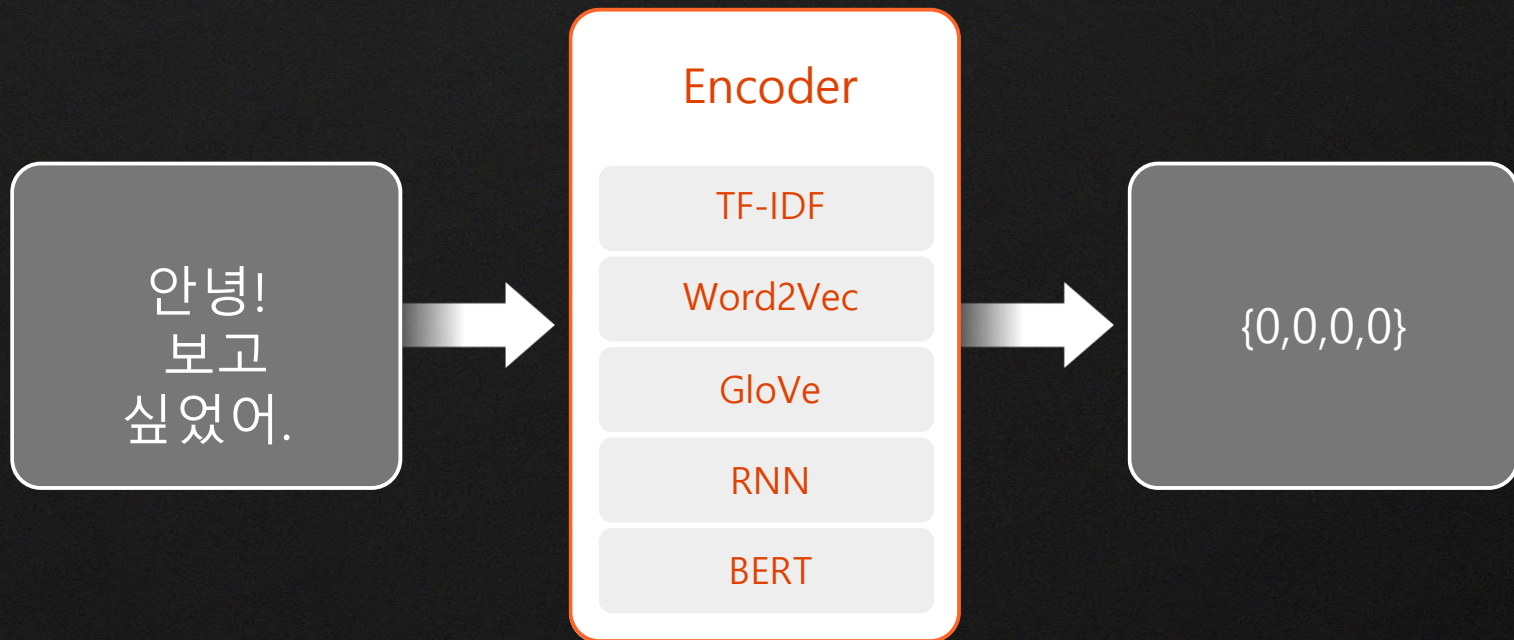
Response DB 생성



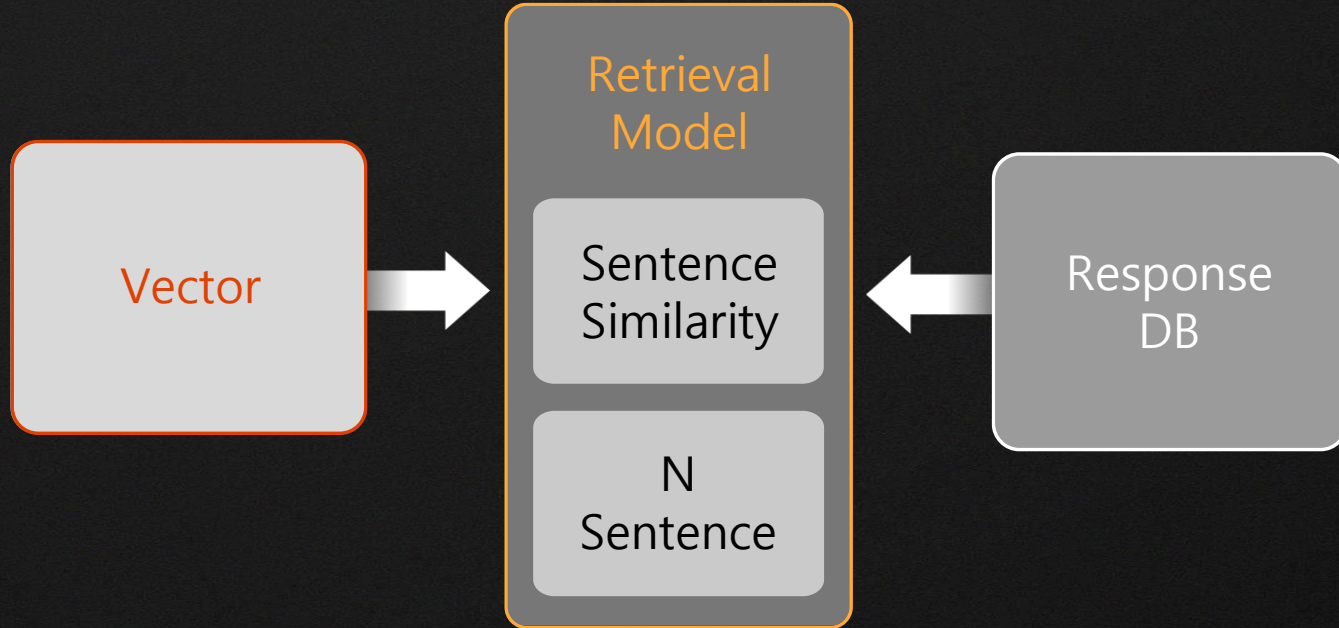
Response DB 생성 예시



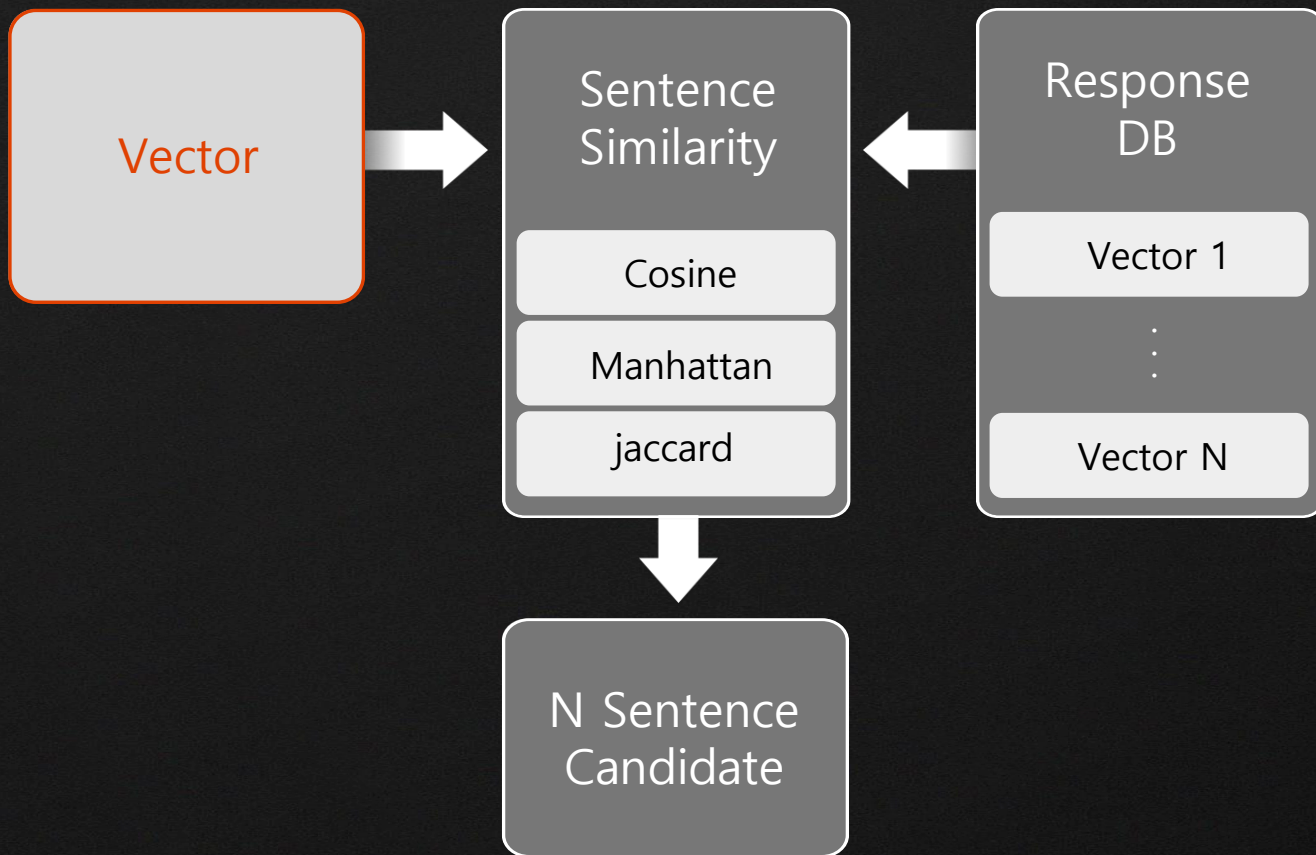
Encoder 예시



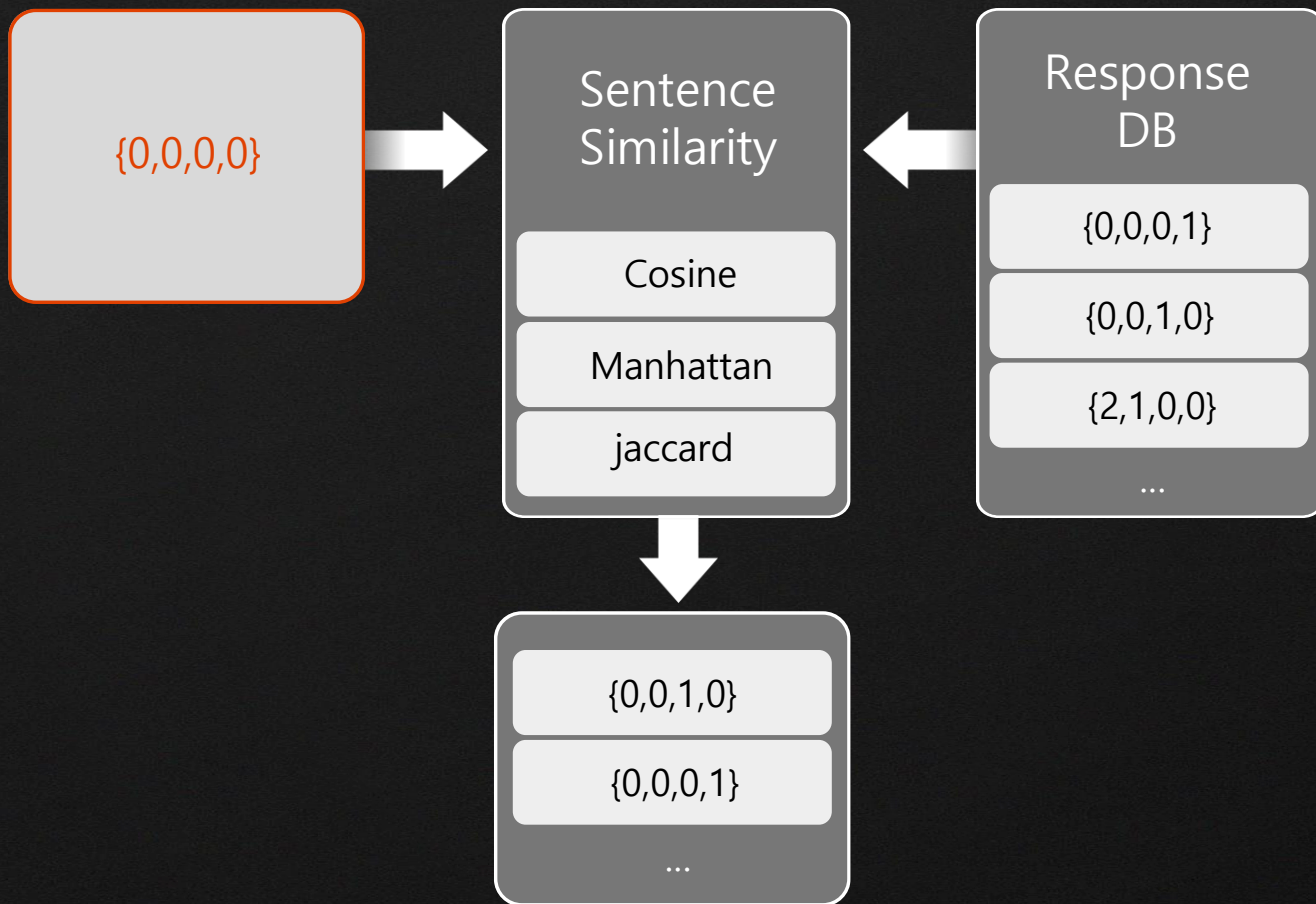
Retrieval Model



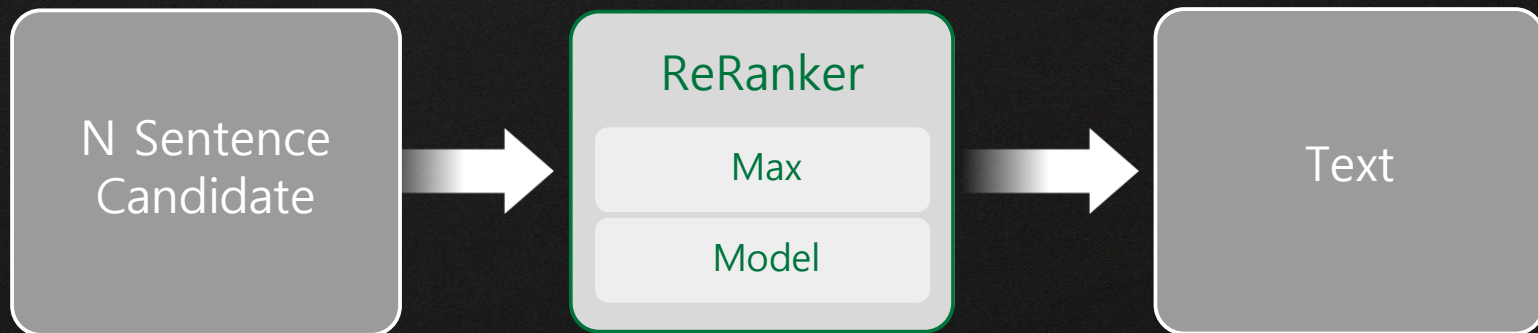
Retrieval Model



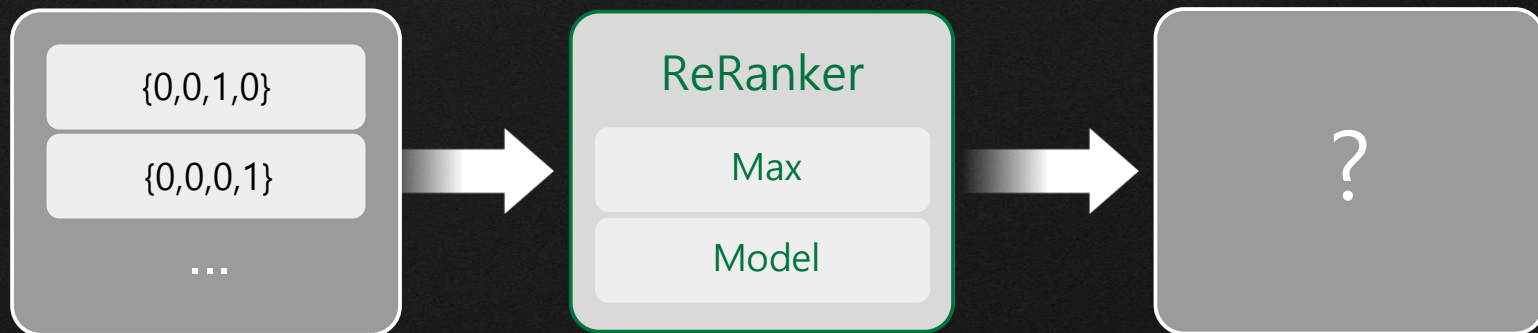
Retrieval Model 예시



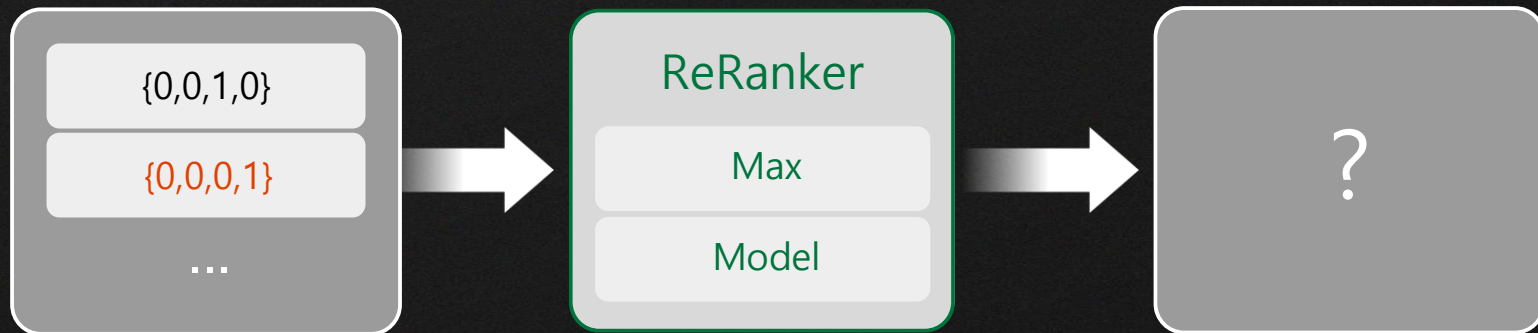
ReRanker



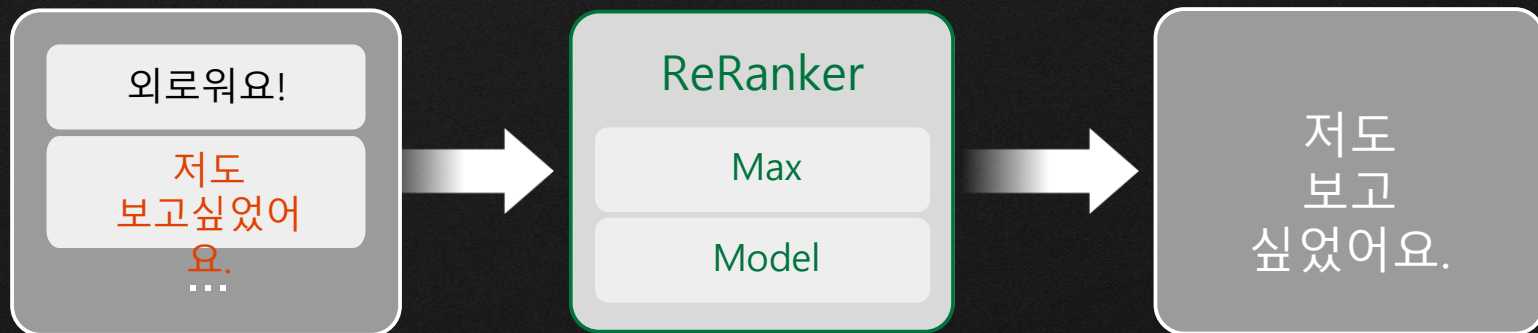
ReRanker 예시



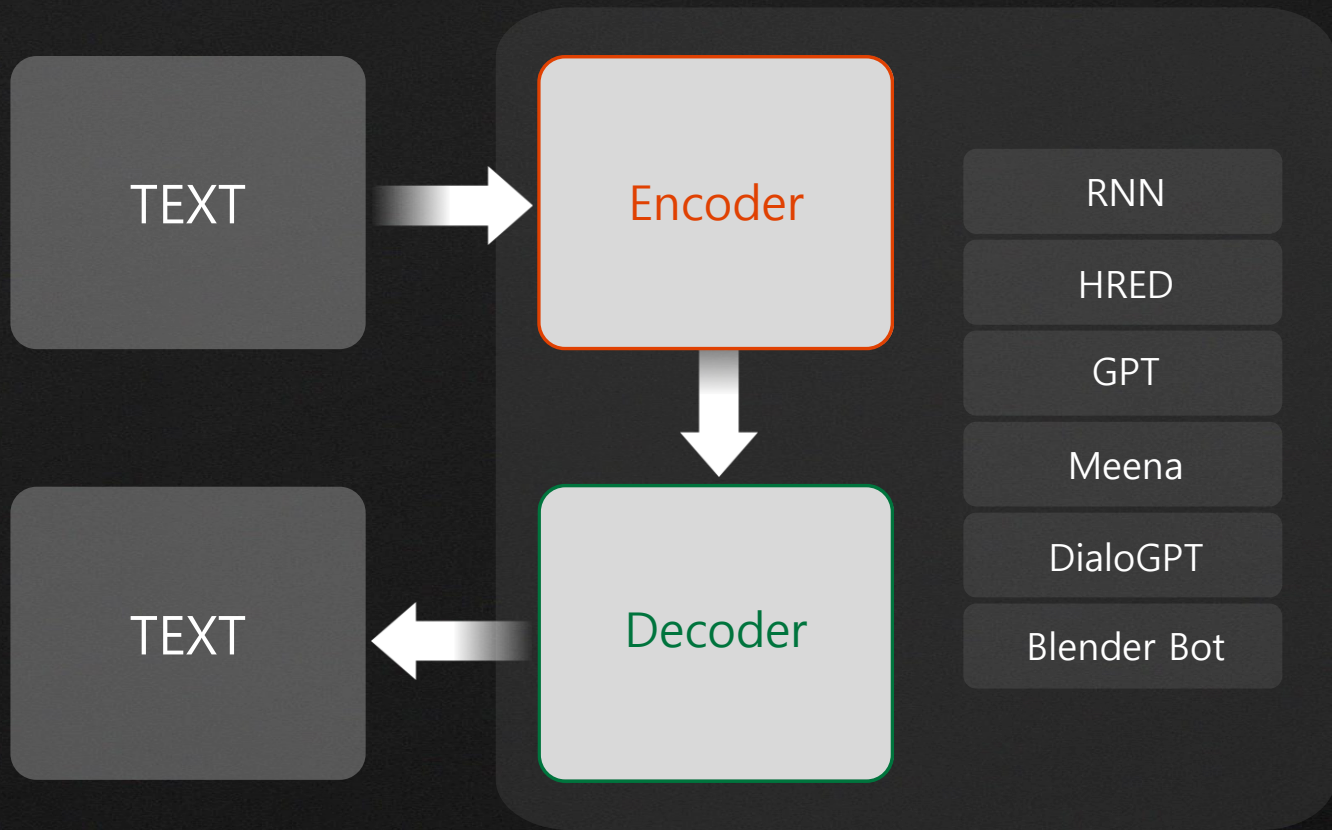
ReRanker 예시



ReRanker 예시



생성기반 방식



생성기반 방식 예시

