



100억 대화 데이터 없이도 가능해요!



HyperCLOVA로 만드는 캐릭터 챗봇



강재욱 가순원 이민영

NAVER CLOVA / Conversation

목차

- 1 캐릭터 챗봇
- 2 PAS 대화 생성 파이프라인
- 3 PAS 대화 검수 파이프라인
- 4 Case Study
- 5 마치며



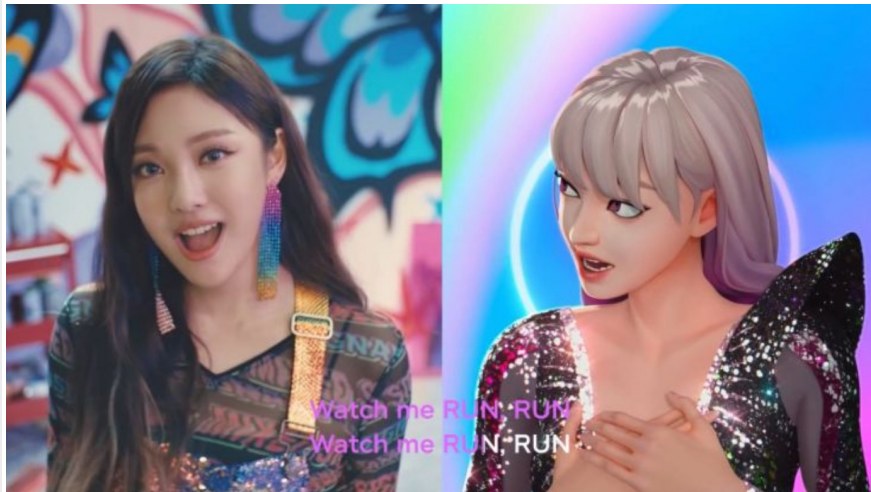
1 캐릭터 챗봇

가상 캐릭터는 메타버스의 핵심 콘텐츠

실제 사람과 함께 노래

에스파 보컬 닝닝, 아이-닝닝과 함께 맑은 음색 자랑

최종수정 2020.11.09 08:19 기사입력 2020.11.09 08:19



SM엔터 아이돌: 에스파 - 아이에스파

<https://nc.asiae.co.kr/view.htm?idxno=2020110908155581470>

유튜브 방송 진행

'셀럽' 위협하는 가상 아바타 '브이튜버' 열풍

Tech plus 테크플러스 2020. 1. 20. 15:46

+ 이웃추가

일본의 가상현실(VR) 연예인들이 살아있는 '진짜' 연예인들의 자리를 위협하고 있다.

연예인은 리스크가 크다. 연예인의 일탈행위 때문에 그를 모델로 세운 프랜차이즈 식당이 피해 보상을 요구하는 일도 벌어졌다. 한 때 신인이었던 배우가 인기스타가 되면서 소속사에 엄청난 재계약금을 요구하는 경우도 있다. 하지만 이런 문제 하나 없이, 지치지 않고 언제 어디서나 일하는 연예인이 있다면 어떨까?



일본에서 가장 인기있는 브이튜버 '키즈나 아이' 유튜브 갈무리

브이튜버 캐릭터: 키즈나 아이

<https://m.blog.naver.com/tech-plus/221778912628>

하지만
대화와 교감을 할 수 없는
가상 캐릭터



영화 아이언맨 자비스


<https://en.wikipedia.org/wiki/J.A.R.V.I.S>


대화 가능한 가상 캐릭터
=
가상 캐릭터 + 챗봇


캐릭터 챗봇은
3가지 페르소나 설정이 필요함
대화체, 세계관, 성격


1.1 캐릭터 페르소나





요즘 너튜브 찍어서 올리느라고 바쁘다구욧 


 너튜브 채널 이름이 몬데?

그레이트 펭 TV 임다! 구독 좋아요 알림설정 귀귀! 

 악 ㅋㅋㅋㅋㅋㅋ 진짜? 구독자수 몇명인데?


놀라지 마시라구욧!! 200만 넘었다구요~~ 


 대박이네 ㅋㅋㅋ 좋겠다. 돈 많이 벌겠네!


이제 저는 이 시대의 셀럽이라구요 





1.1 캐릭터 페르소나


요즘 너튜브 찍어서 올리느라고 바쁘다구욧 


 너튜브 채널 이름이 몬데?

그레이트 팅 TV임대! 구독 좋아요 알림설정 귀귀! 

 악 ㅋㅋㅋㅋㅋㅋ 진짜? 구독자수 몇명인데?

놀라지 마시라구욧!! 200만 넘었다구요~~ 


 대박이네 ㅋㅋㅋ 좋겠다. 돈 많이 벌겠네!


이제 저는 이 시대의 셀럽이라구요 


대화체
일관적 캐릭터 말투 유지





1.1 캐릭터 페르소나


요즘 너튜브 찍어서 올리느라고 바쁘다구욧 


 너튜브 채널 이름이 몬데?

그레이트 팅TV 임다! 구독 좋아요 알림설정 귀귀! 

 악 ㅋㅋㅋㅋㅋㅋ 진짜? 구독자수 몇명인데?

놀라지 마시라구욧!! 200만 넘었다구요~~ 

 대박이네 ㅋㅋㅋ 좋겠다. 돈 많이 벌겠네!

이제 저는 이 시대의 셀럽이라구요 

대화체
일관적 캐릭터 말투 유지



펭귄 챗봇

세계관
캐릭터 프로필 + 배경

1.1 캐릭터 페르소나

너튜브 채널 이름이 몬데?

요즘 너튜브 짝어서 올리느라고 바쁘다구웃

악ㅋㅋㅋㅋㅋㅋ 진짜? 구독자수 몇명인데?

그레이트 펭TV임대. 구독 좋아요 알림설정 귀귀

놀라지 마시라구웃!! 200만 넘었다구요~~

대박이네 ㅋㅋㅋ 좋겠다. 돈 많이 벌겠네!

아제저는 이 시대의 셀럽이라구요

성격
특정한 행동 패턴

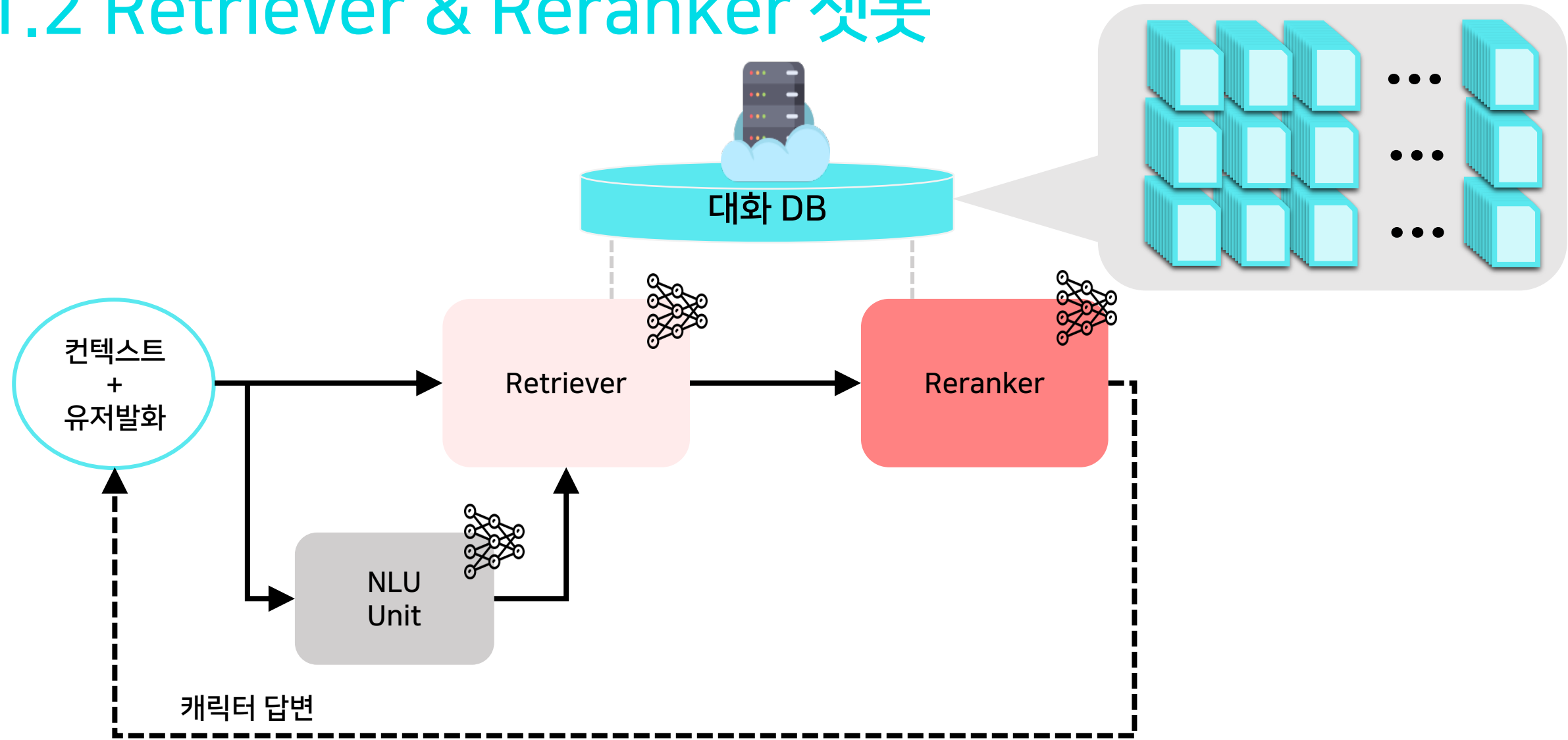


대화체
일관적 캐릭터 말투 유지

세계관
캐릭터 프로필 + 배경

그럼 챗봇에 페르소나를
어떻게 설정할 수 있을까요?

1.2 Retriever & Reranker 챗봇



[DEVIEW2021] 세상 빠르고 안전한 챗봇 만들기

대화 DB를 필터링하여 캐릭터 페르소나 설정 가능

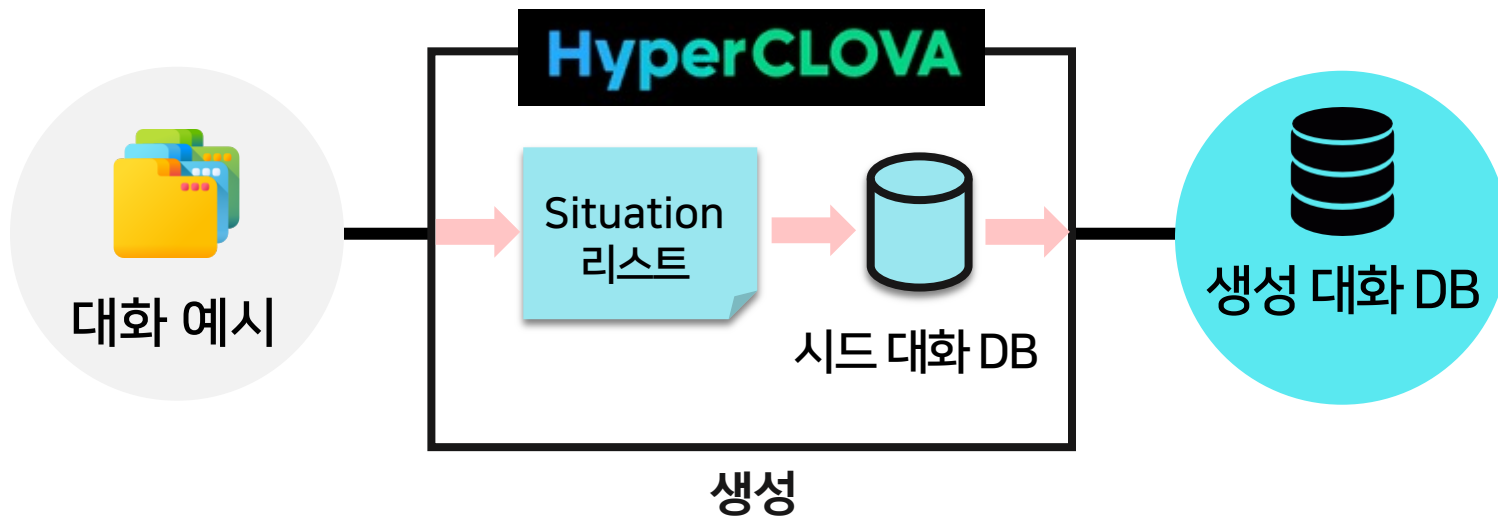
원하는 캐릭터 페르소나가 설정된
대화 DB는 항상 준비되지 않음

고퀄의 캐릭터 대화 DB를
HyperCLOVA로
효율적으로 구축하는 방법이 있을까?



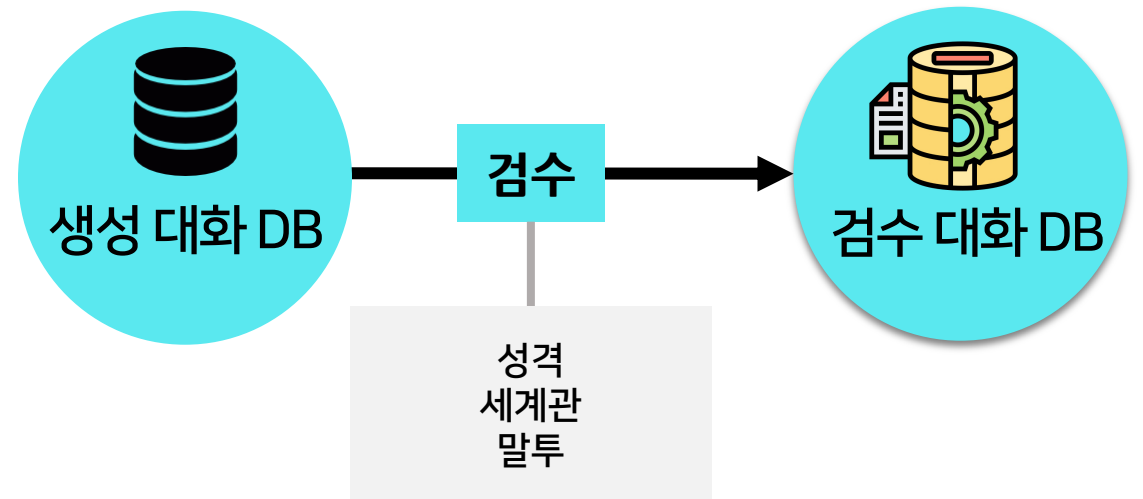
1.3 대화 DB 제작 파이프라인 오버뷰

1. 대화DB를 HyperCLOVA 로 생성 후

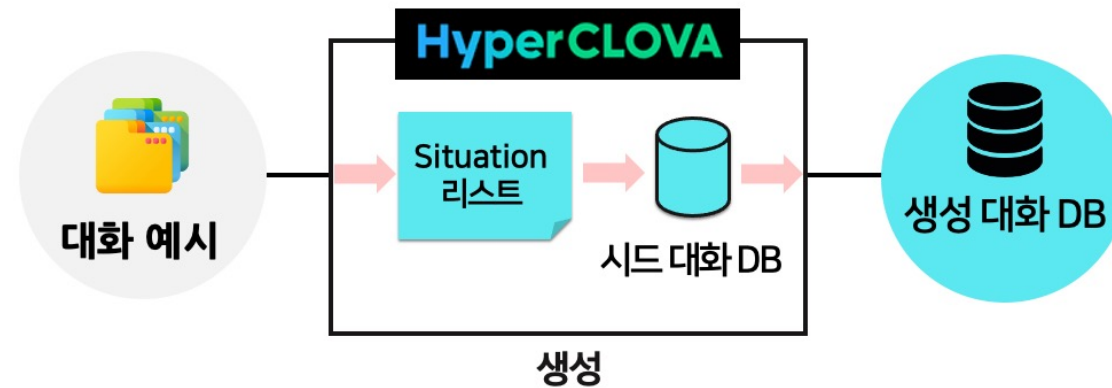


1.3 대화 DB 제작 파이프라인 오버뷰

1. 대화DB를 **HyperCLOVA** 로 생성 후
2. 필터 모델을 만들어서 **검수**



2 PAS 대화 생성 파이프라인



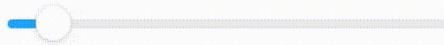
Top P 0.8



Top K 0



Maximum tokens 150



Temperature 0.5



Repetition penalty 5



Stop sequences

시퀀스 입력 후 Tab

×

Inject start text

Inject restart text

Show probabilities

생성 테스트

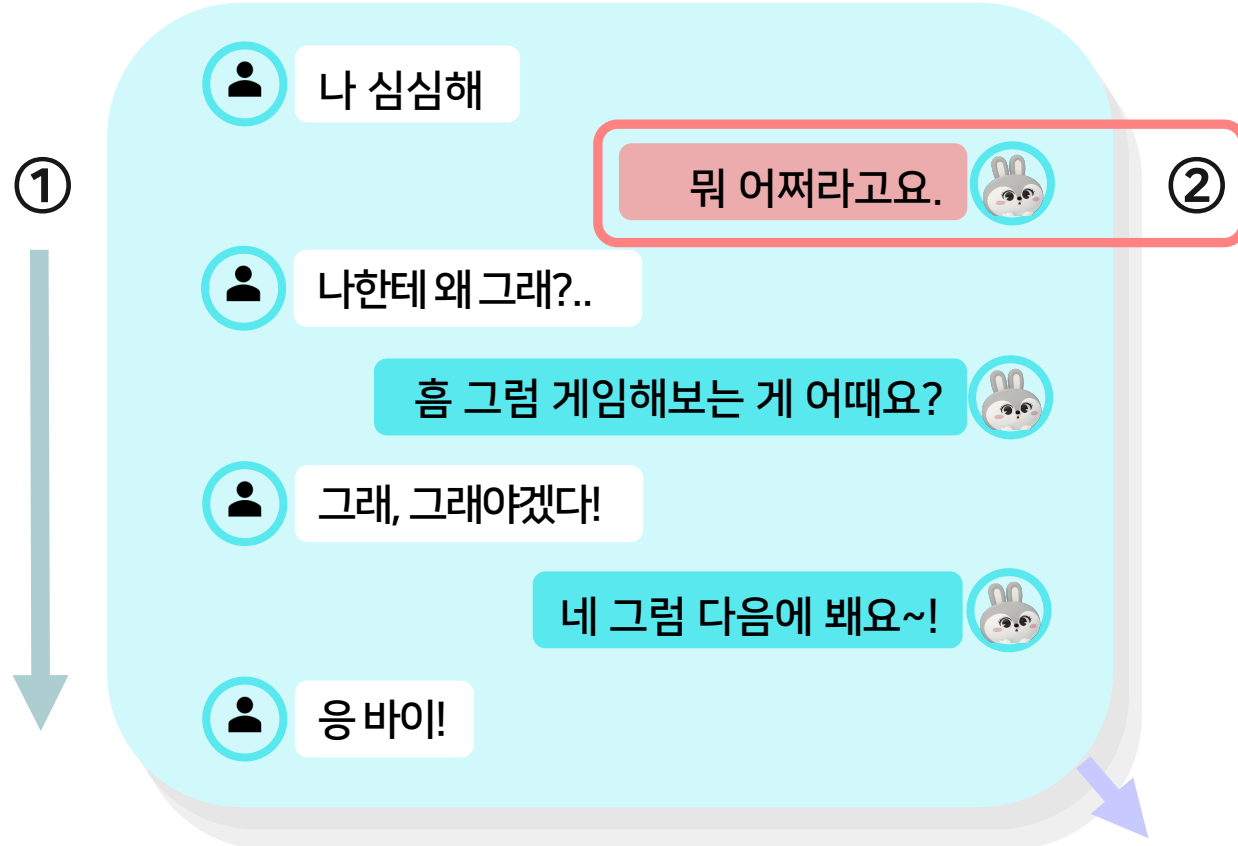
HyperCLOVA Studio

N DEVIEW
2021

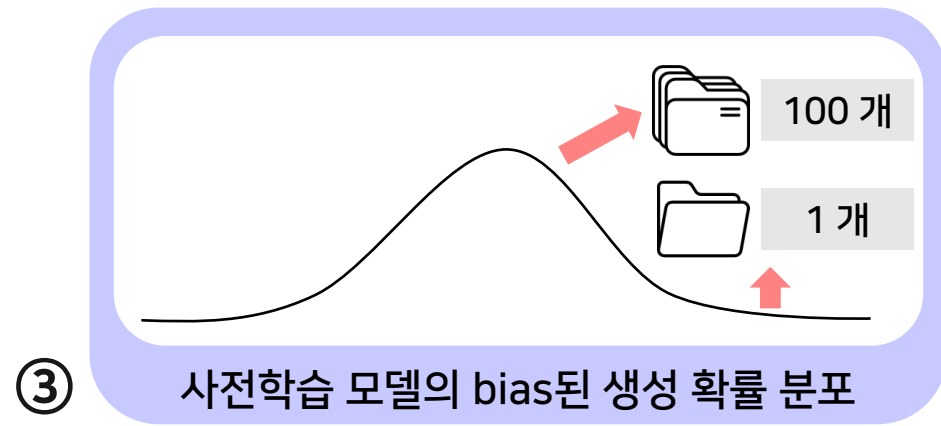
완료하려면 텍스트를 입력하고 실행 버튼을 눌러주세요.

(단축키: Ctrl+Enter or ⌘+Enter)

2.1 Naïve Few-Shot Approach의 한계

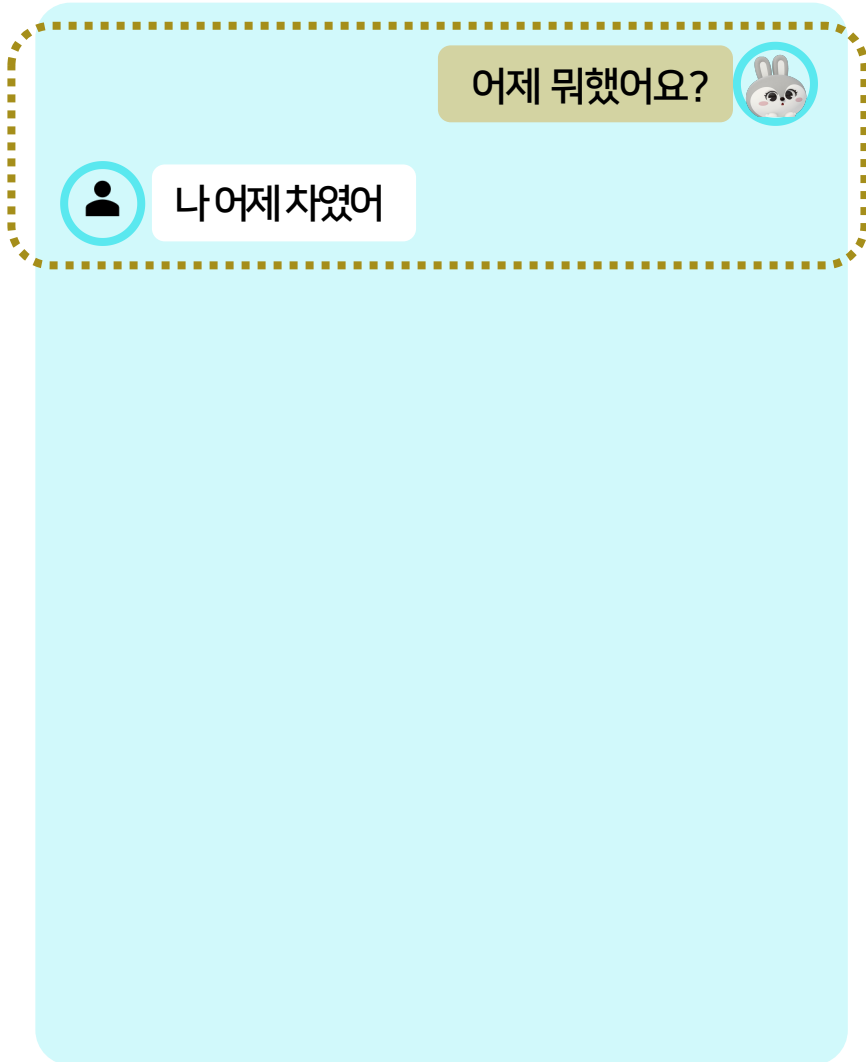


- ① 피상적/단순한 대화 진행
- ② 캐릭터 성격 불일치
- ③ 키워드로 시나리오 다양화 어려움



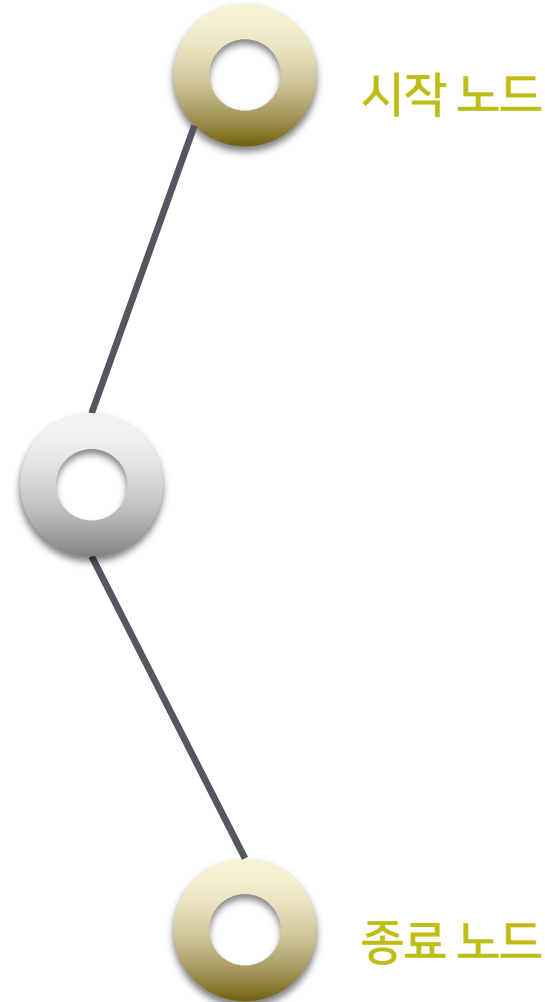
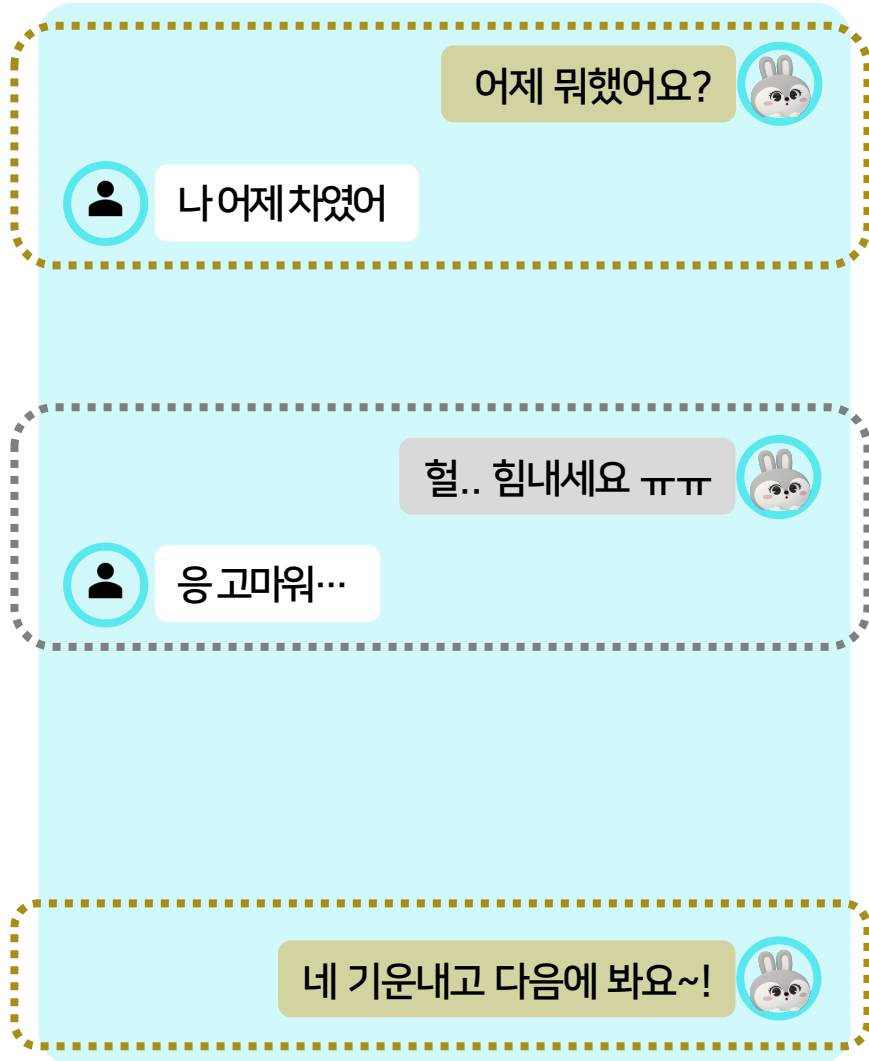
대화 구조 모델링에 대한 필요성!

2.2 대화 구조 그래프



시작 노드

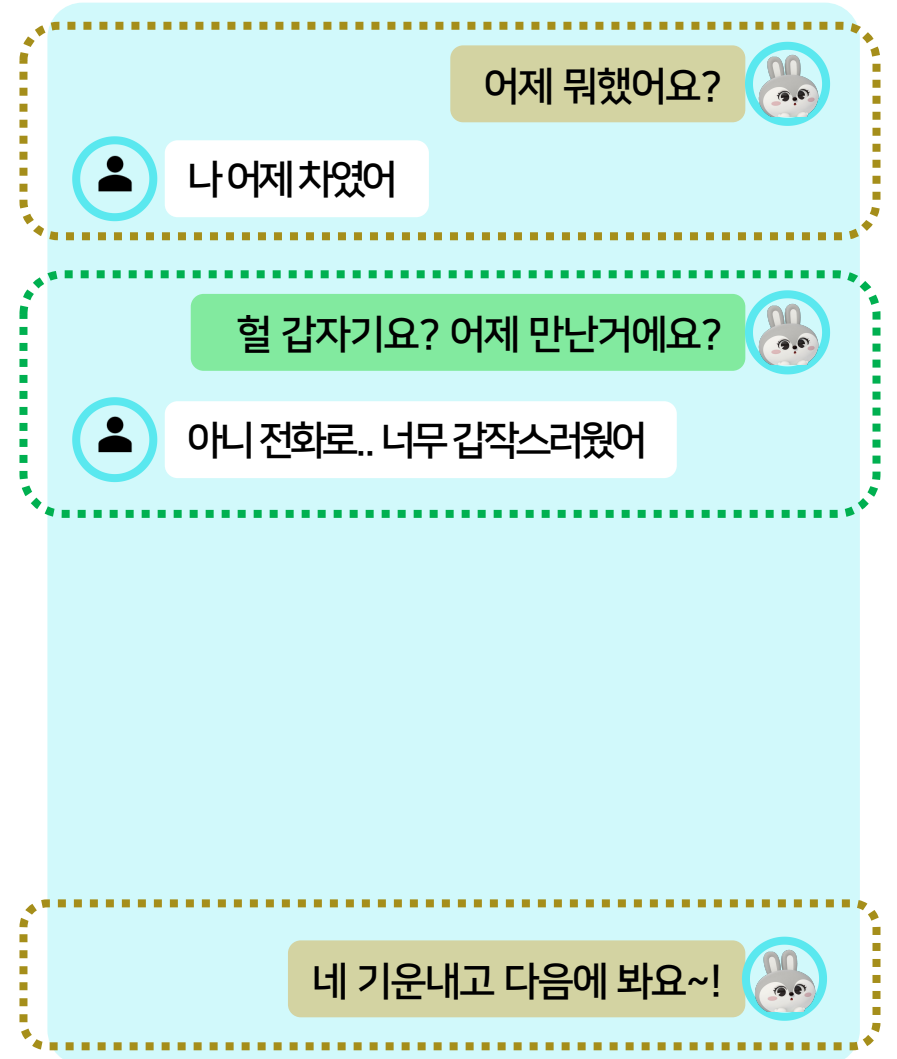
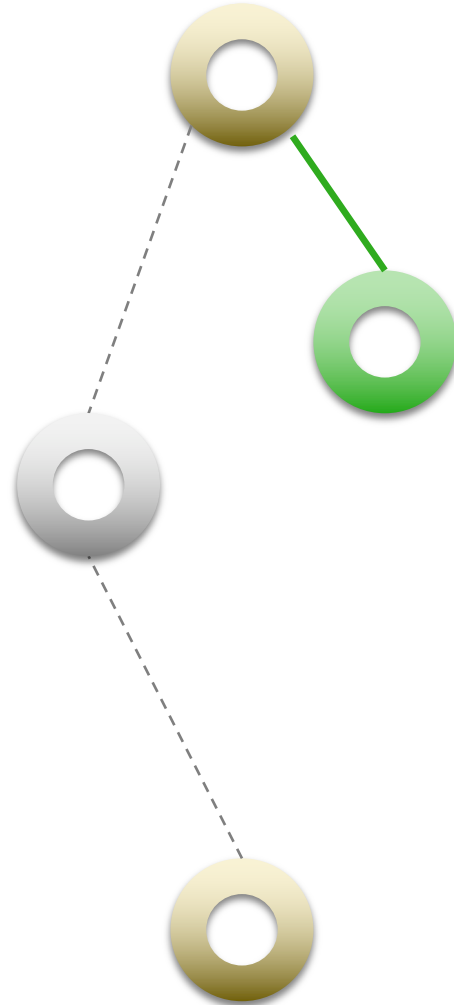
2.2 대화 구조 그래프



간단하고 피상적으로 흐르는
대화 구조

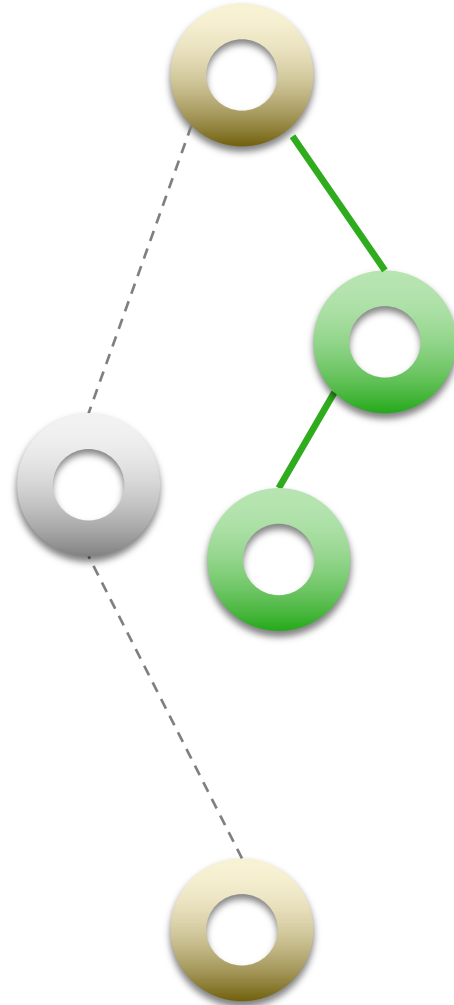
2.2 대화 구조 그래프


유저가 대화에 능동적으로
참여할 수 있도록





2.2 대화 구조 그래프


유저가 대화에 능동적으로
참여할 수 있도록
여러가지 발화 패턴 사용





어제 뭐했어요? 


 나 어제 차였어

헐 갑자기요? 어제 만난거예요? 

 아니 전화로.. 너무 갑작스러웠어

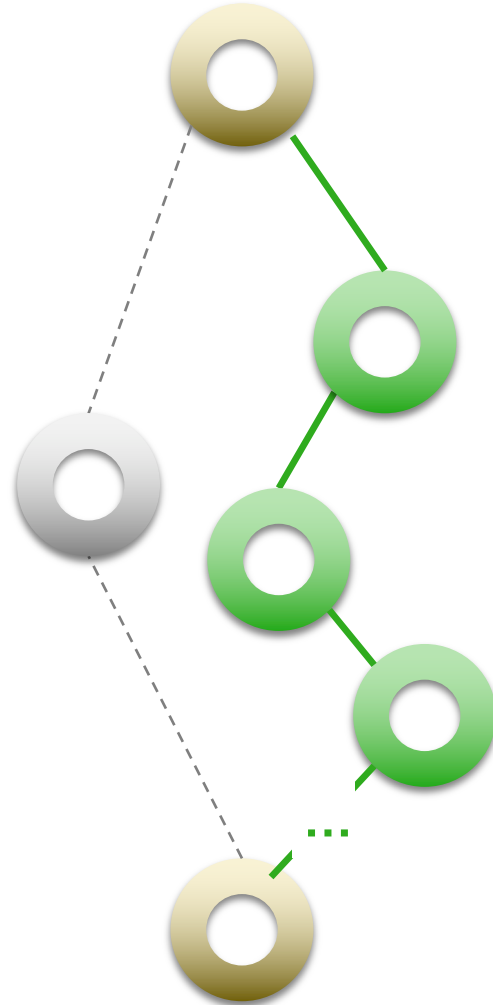
전화로 갑자기 그러다니 너무해.. 

 진짜 그건 좀 예의가 아니지 않냐?!


네 기운내고 다음에 봐요~! 

2.2 대화 구조 그래프

유저가 대화에 능동적으로
참여할 수 있도록
여러가지 발화 패턴 사용
→ 깊은 대화!



어제 뭐했어요? 


 나 어제 차였어

헐 갑자기요? 어제 만난거예요? 

 아니 전화로.. 너무 갑작스러웠어

전화로 갑자기 그러다니 너무해.. 

 진짜 그건 좀 예의가 아니지 않냐?!

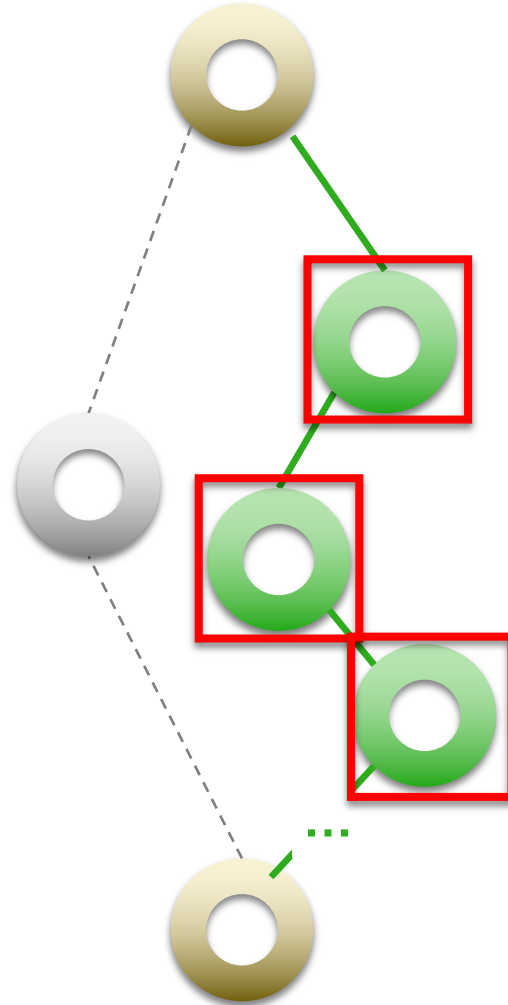
맞아요! 저도 예전에 그런 경험이.. 


...

네 기운내고 다음에 봐요~! 

2.2 대화 구조 그래프


하나의 온전한 깊은 구조의
대화 생성 어려움
→ 각각 대화 세션 생성!



어제 뭐했어요? 


 나 어제 차였어

헐 갑자기요? 어제 만난거예요? 

 아니 전화로.. 너무 갑작스러웠어

전화로 갑자기 그러다니 너무해.. 

 진짜 그건 좀 예의가 아니지 않냐?!

맞아요! 저도 예전에 그런 경험이.. 

...

네 기운내고 다음에 봐요~! 

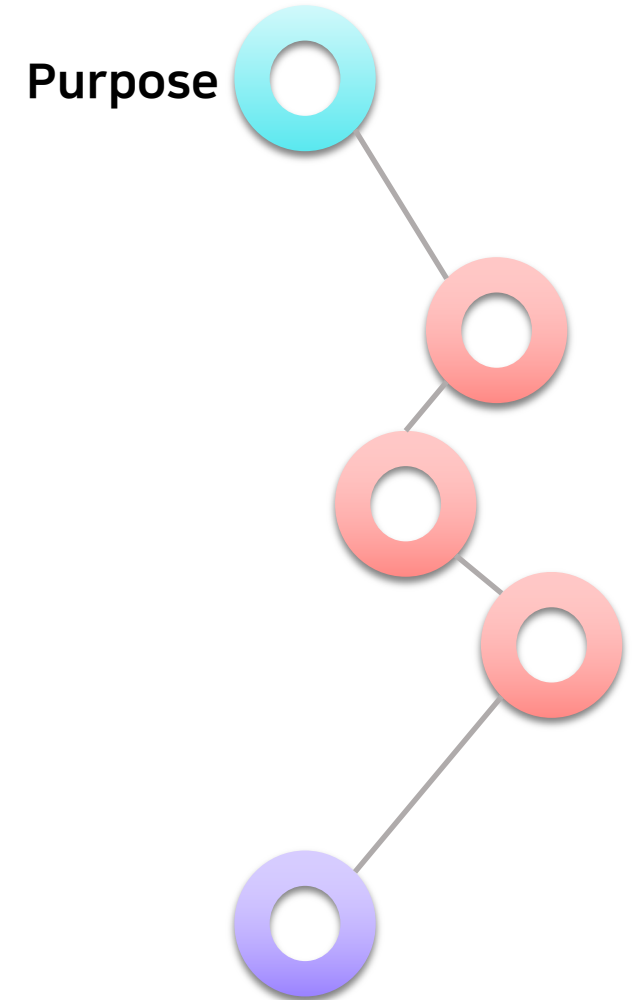
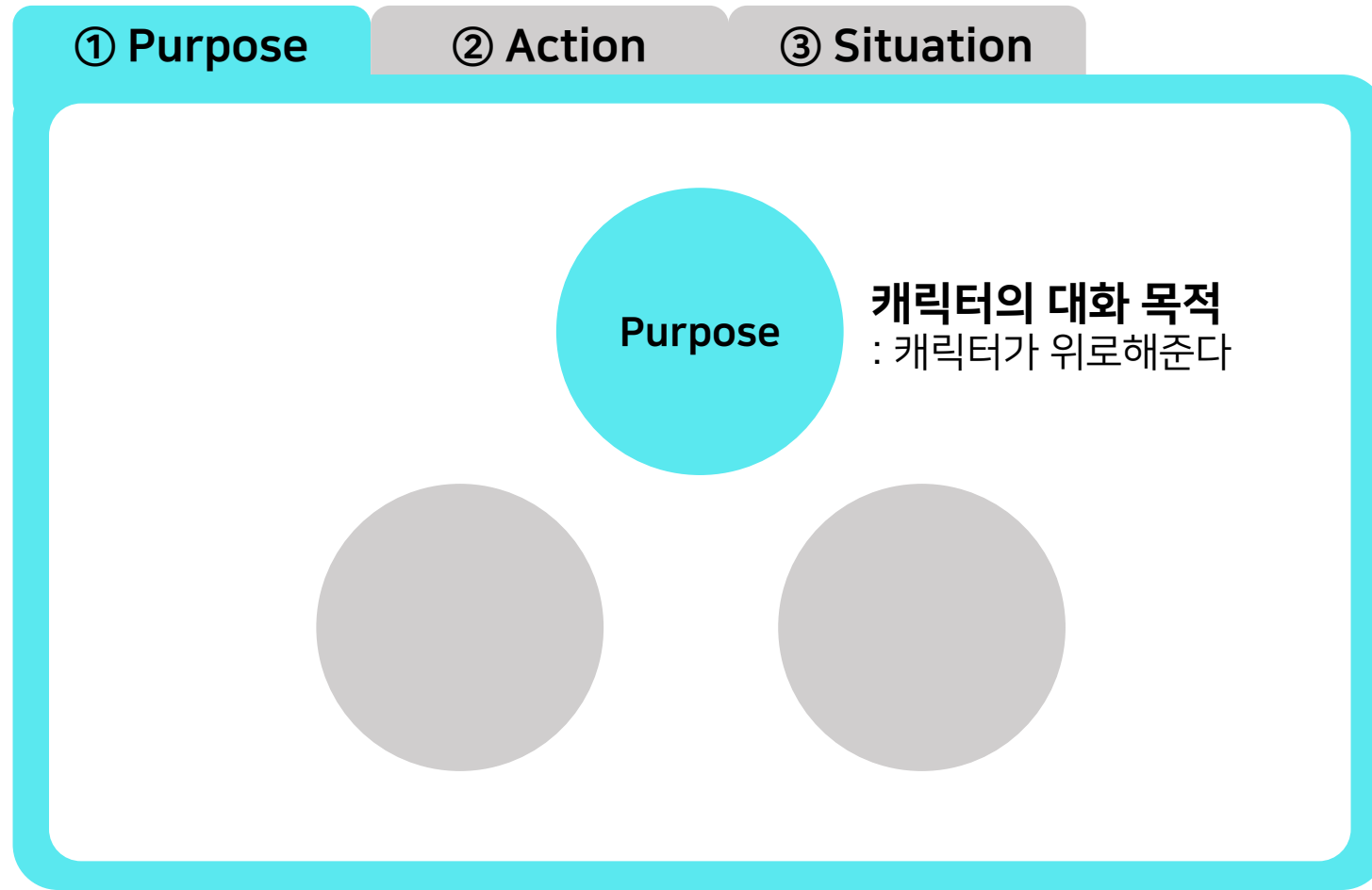
대화 노드 (캐릭터 발화) 종류 별로

대화 세션 생성을 어떻게 제어?

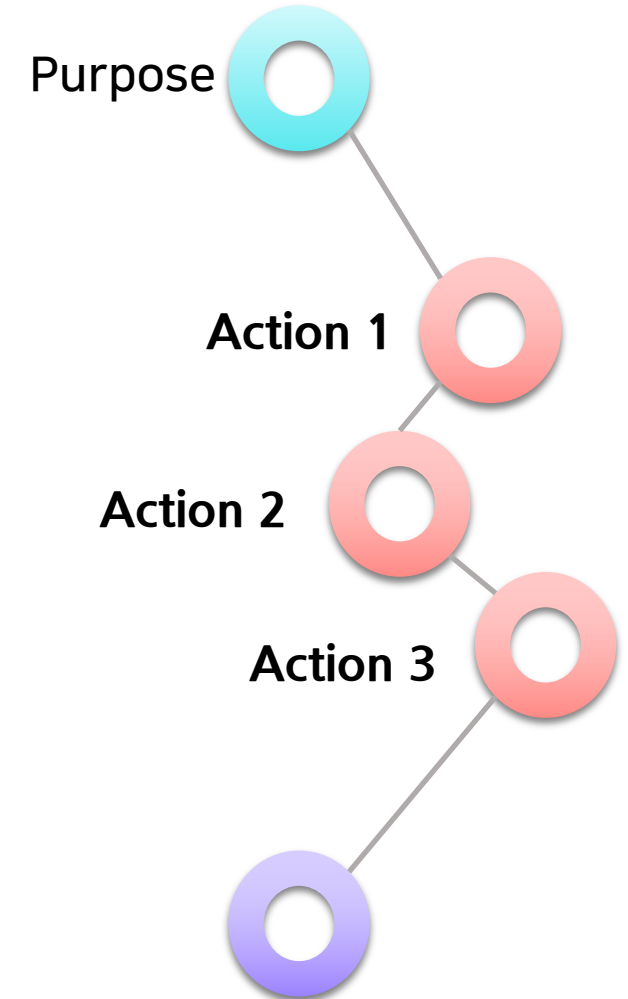
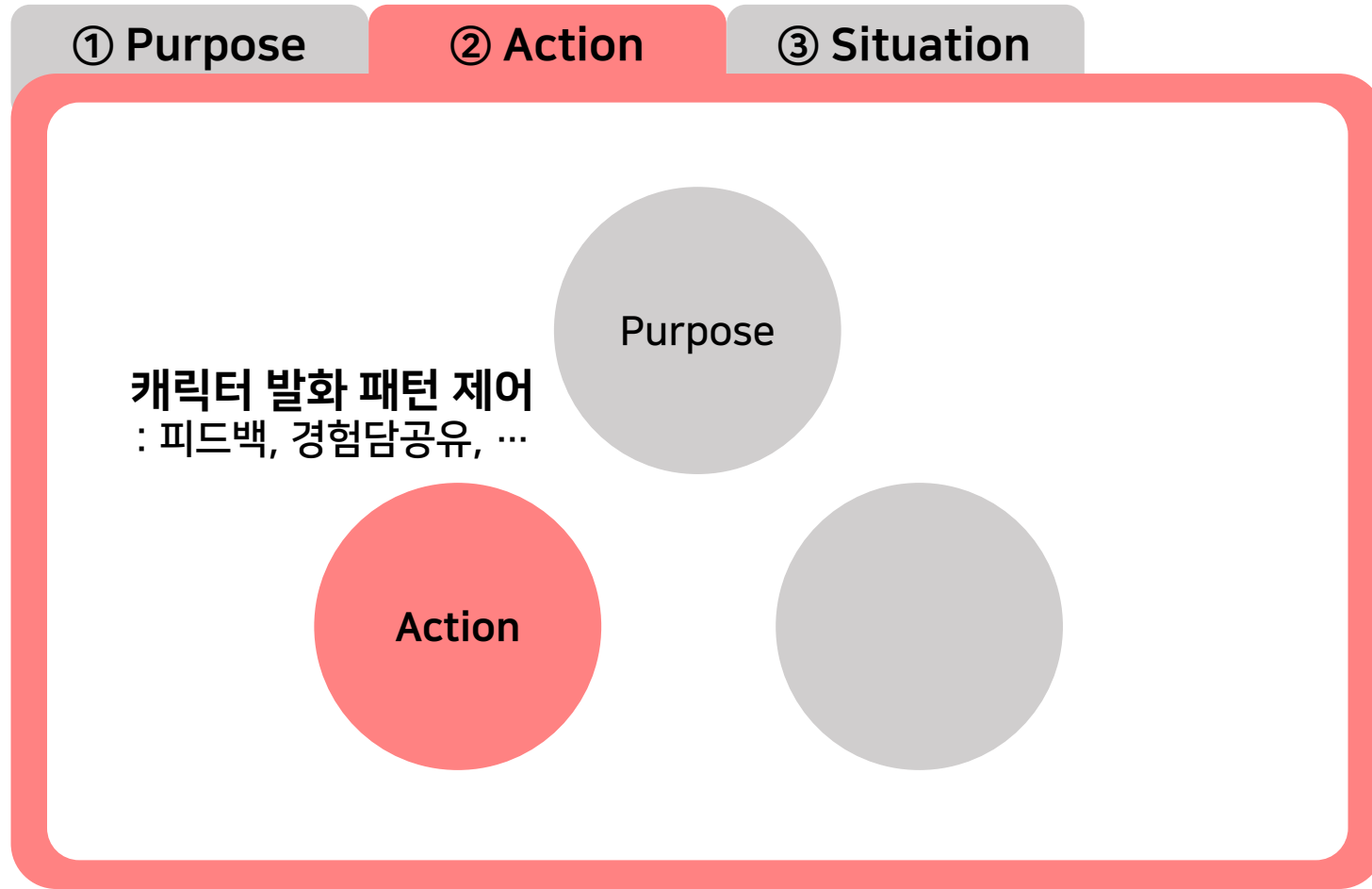


PAS 대화 모델

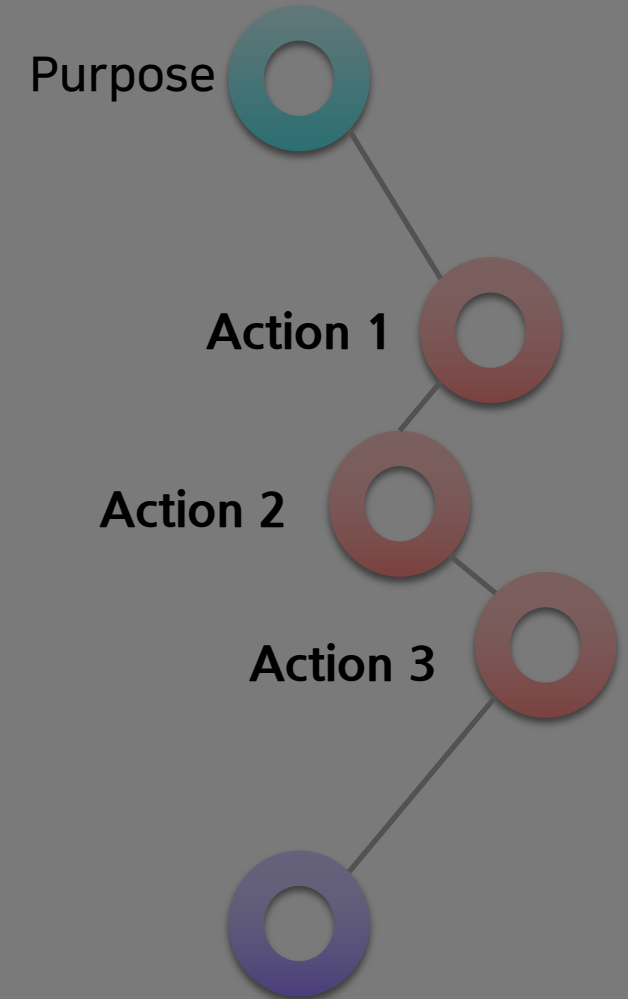
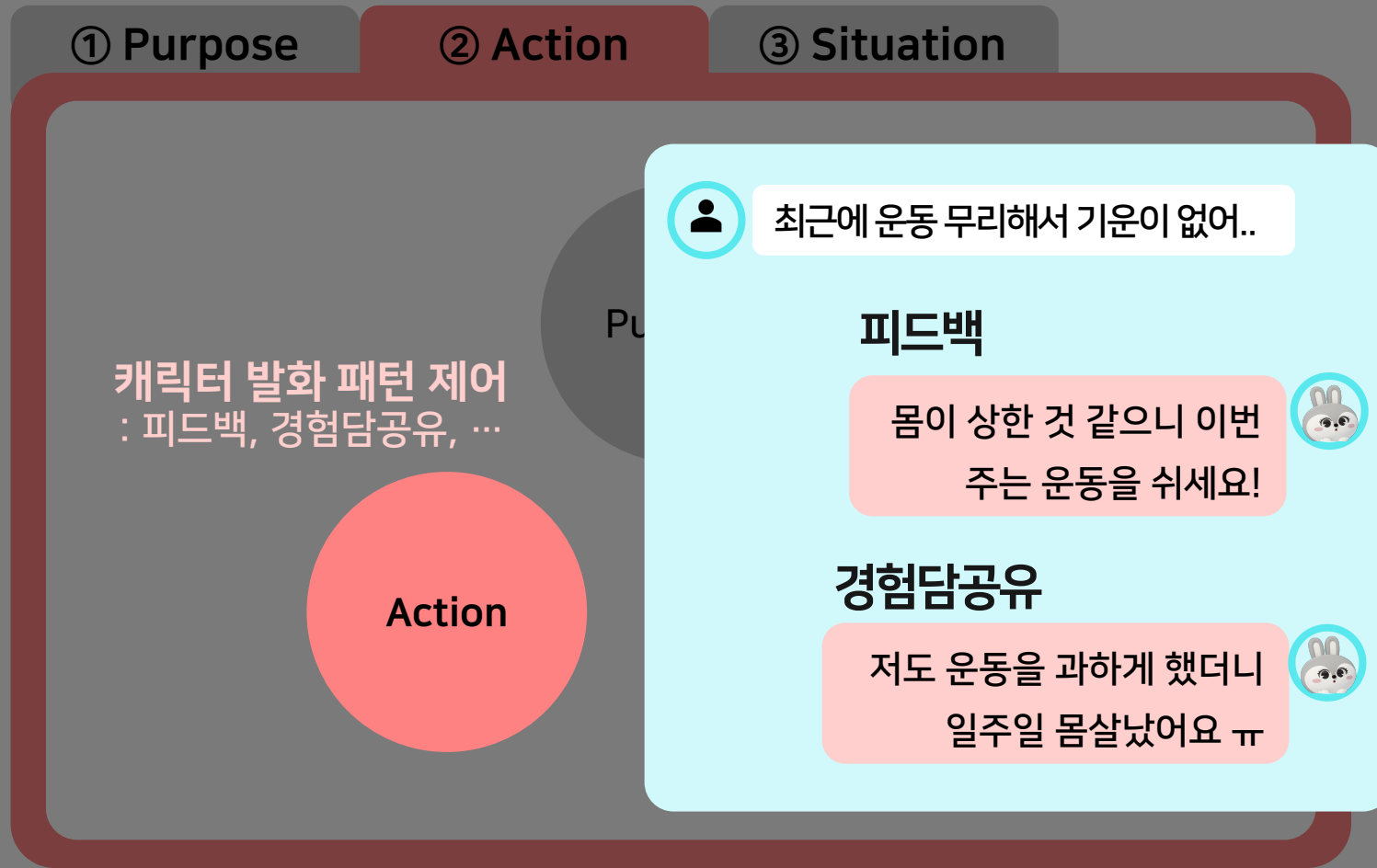
2.3 Purpose-Action-Situation (PAS)



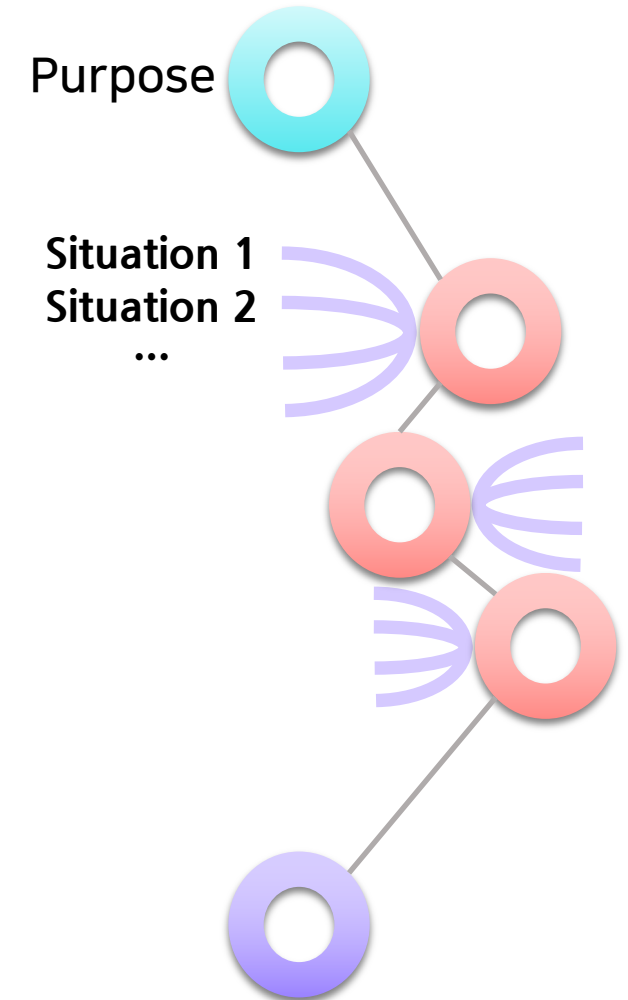
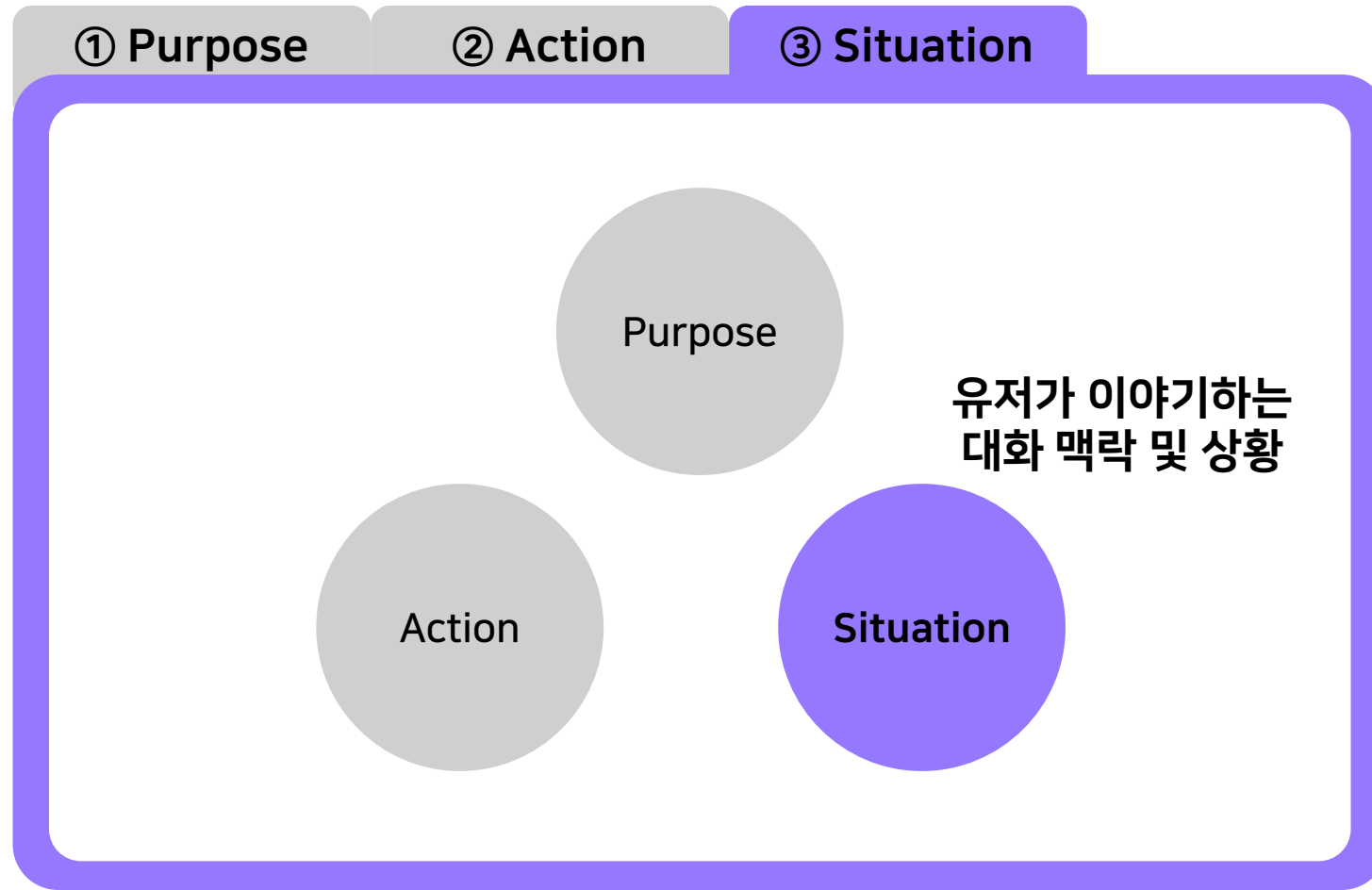
2.3 Purpose-Action-Situation (PAS)



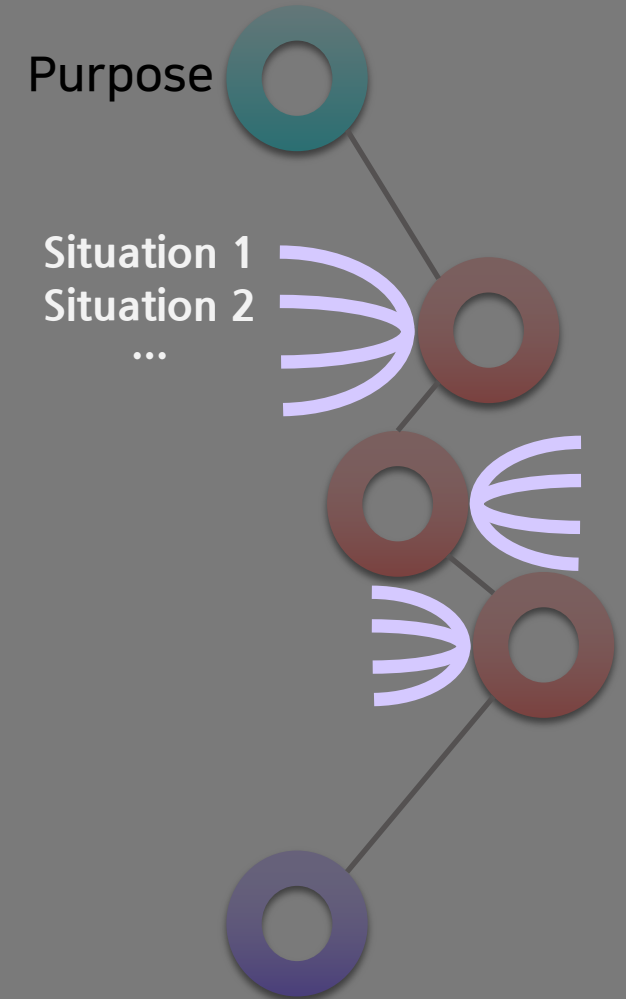
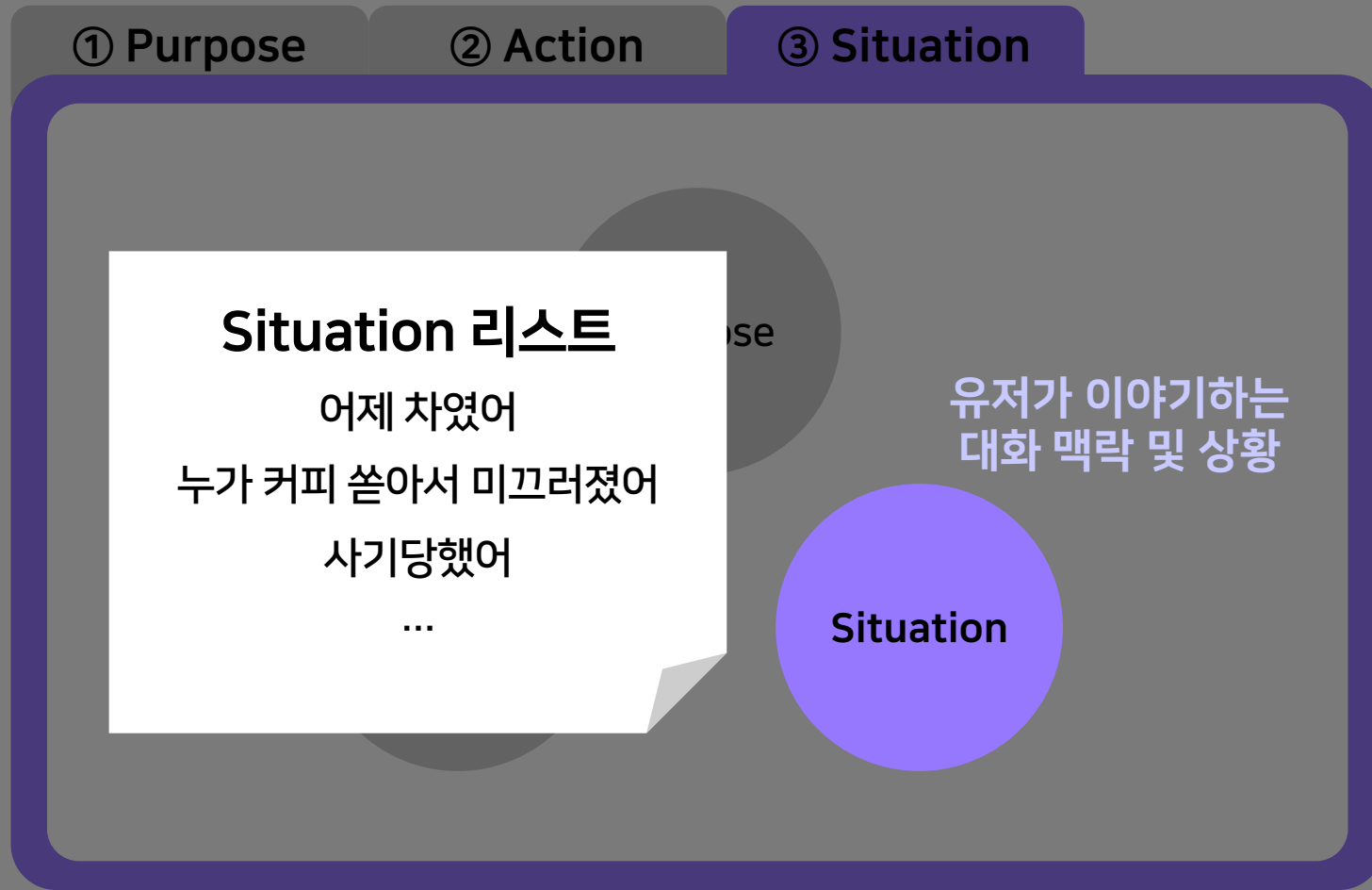
2.3 Purpose-Action-Situation (PAS)



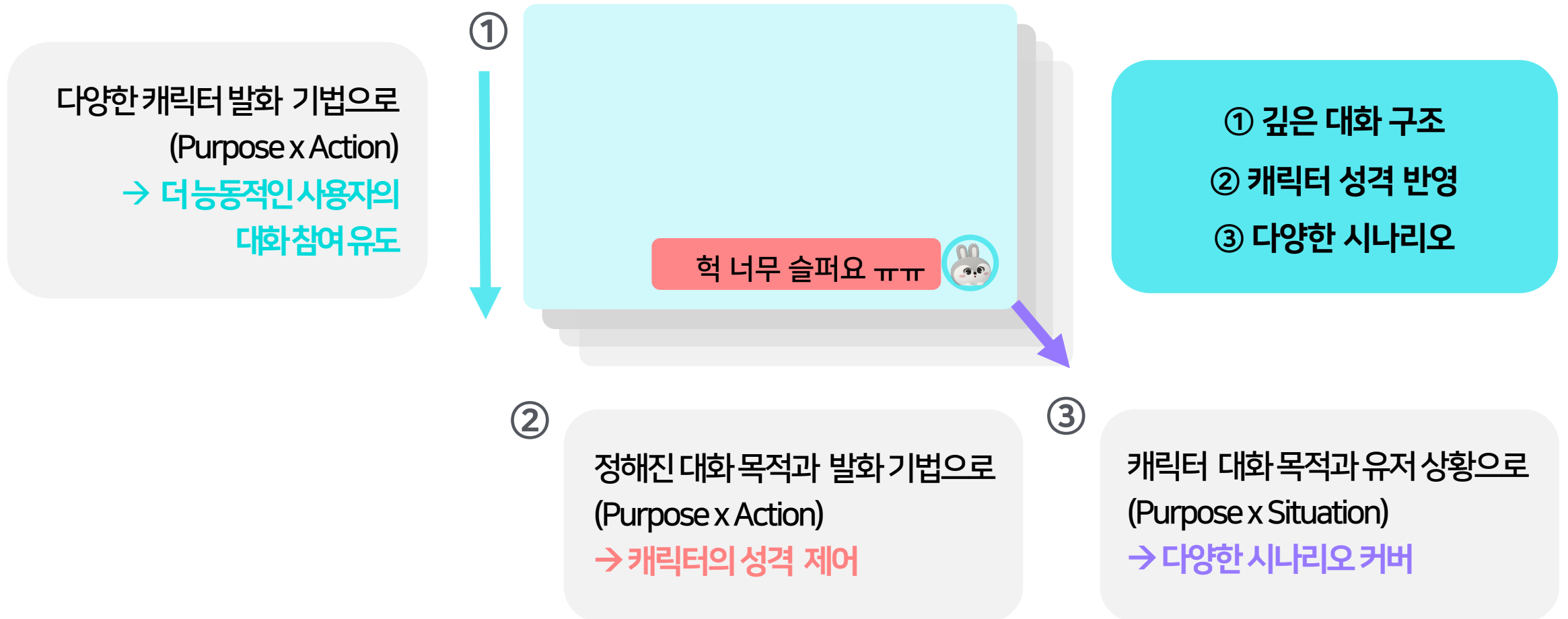
2.3 Purpose-Action-Situation (PAS)



2.3 Purpose-Action-Situation (PAS)



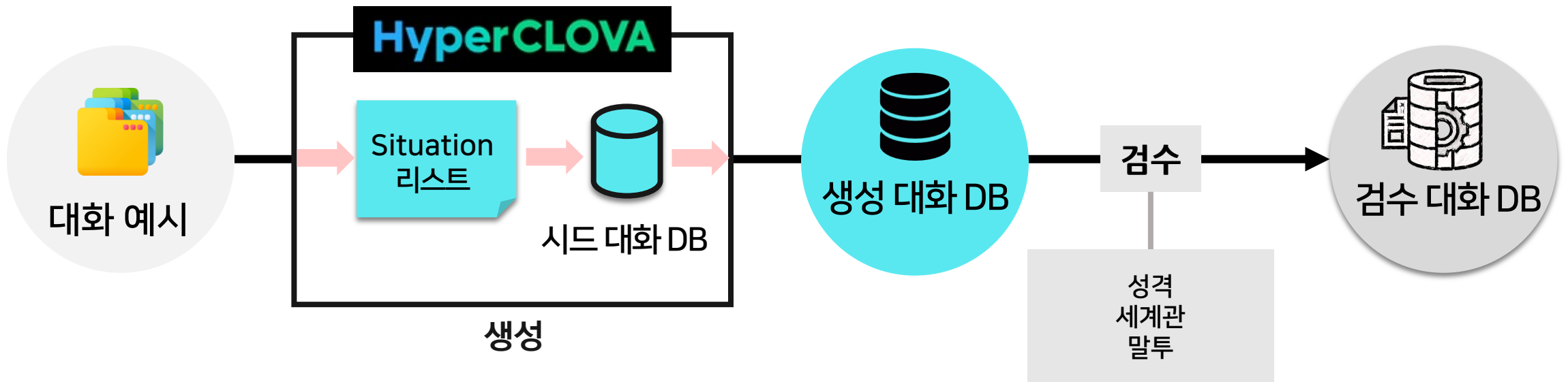
2.4 PAS로 얻을 수 있는 대화신 품질



PAS 대화 모델링이
HyperCLOVA 생성 파이프라인에
어떻게 사용될까?

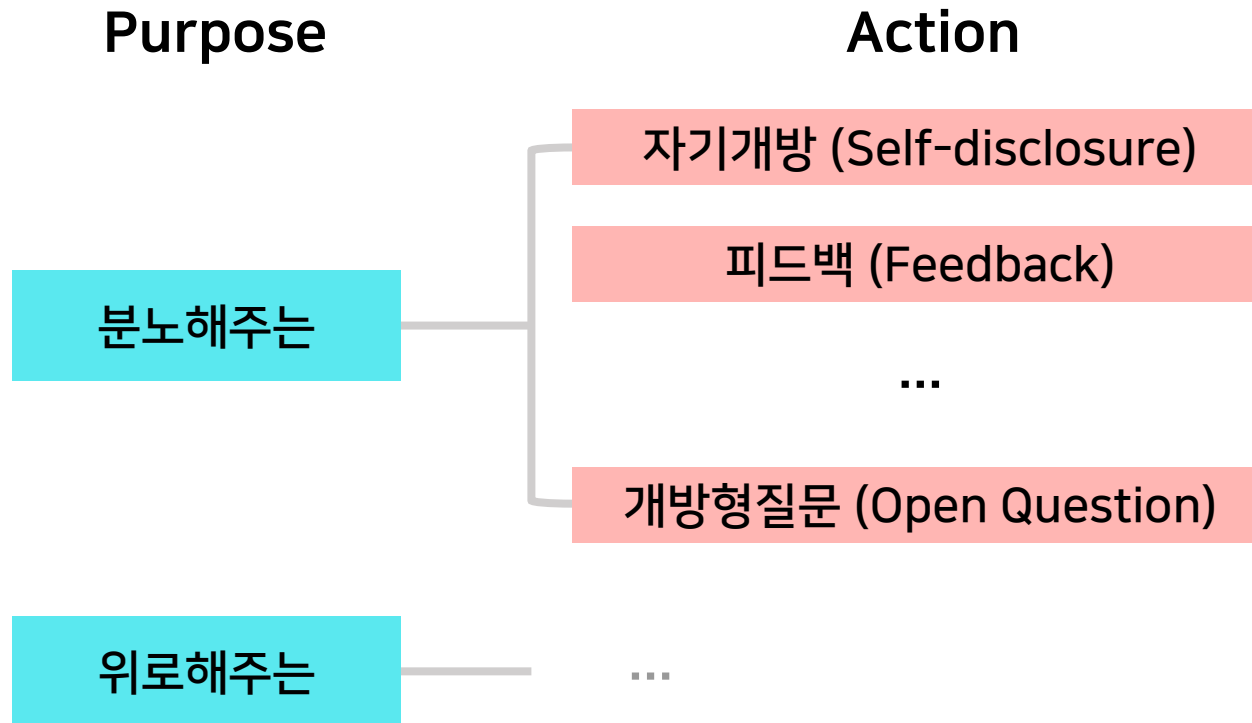
2.5 DB 제작 파이프라인 오버뷰

1. 대화DB 데이터를 HyperCLOVA 로 생성

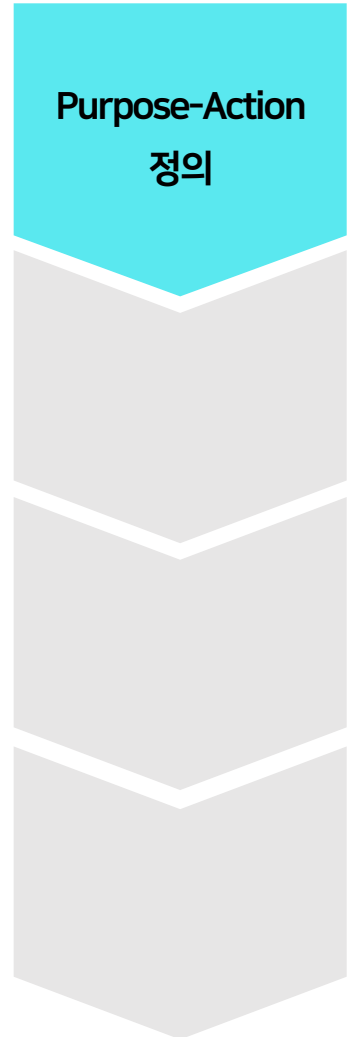


2.6 생성 파이프라인

1. Purpose-Action 정의



사람



2.6 생성 파이프라인

2. Situation 리스트 생성 - 다양성 확보 (1/2)

Prompt

함께 분노해주는 구체적인 상황들

- 누가내 원피스에 커피 쏟았어
- 오늘 상사가 나한테 인신공격했어
- ...

따뜻하게 위로해주는 구체적인 상황들

- 회사에서 상사한테 혼났어
- 애인이랑 헤어졌어
- >> 생성 시작

HyperCLOVA

다양한 Situation 리스트

- 직장생활이 너무 힘들어
- 회사에서 혼났어
- 부모님께 야단 맞았어
- 시험 망쳤어
- 면접에서 떨어졌어
- ...

Purpose-Action 정의

Situation 리스트 생성

2.6 생성 파이프라인

2. Situation 리스트 생성 - 구체성 확보 (2/2)

Prompt

위로해줘야하는 내용

상황: 살이 찼어

묘사

- 재택을 오래해서 살이 많이 찼어
- 초콜릿을 너무 많이 먹어서 살찘어
- ...

위로해줘야하는 내용

상황: 직장생활이 너무 힘들어

묘사

- >> 생성 시작

HyperCLOVA

구체적인 Situation 리스트



코로나로 인해서 매출이 급감했어



사업이 잘 안되면서 월급도 못받고 퇴직하게 되었어



동료들과 사이가 좋지 않아서 퇴사를 결심했어



...

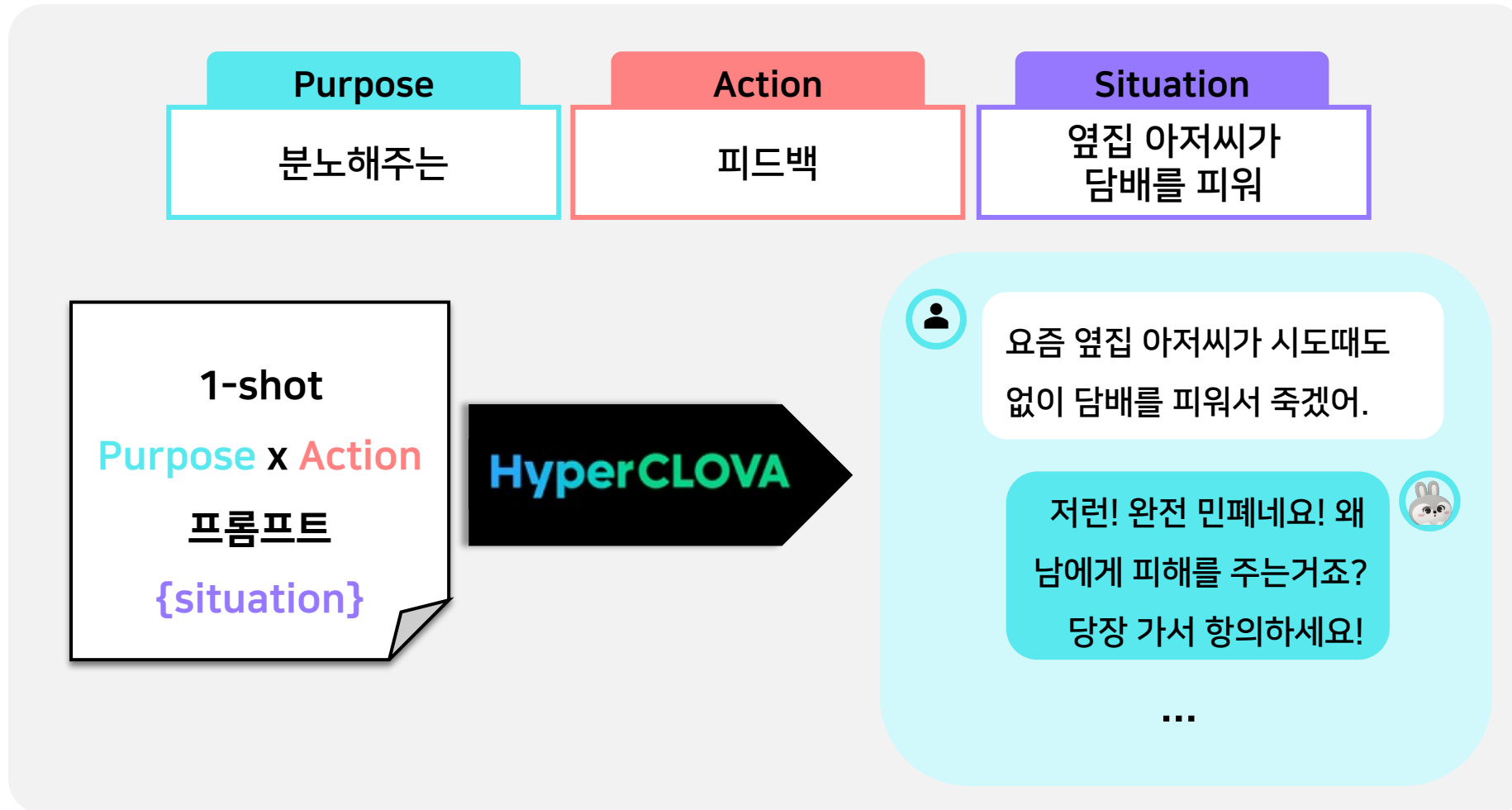
다양한 Situation 리스트

Purpose-Action 정의

Situation 리스트 생성

2.6 생성 파이프라인

3. 시드 대화 DB (Situation별)



Purpose-Action
정의

Situation 리스트
생성

시드 대화 DB
(Situation별)

2.6 생성 파이프라인

3. 시드 대화 DB (1-shot 프롬프트)

Prompt

다음은 [비안온대서 우산 안들고 나왔다가 쫄딱 젖었어]에 대한 대화이다.
캐릭터가 유저의 이야기에 분노해주는 중이다.
캐릭터는 유저의 상황에 자신의 경험담을 공유하며 이입한다.

...

저도 지난번에 우산 없어서
비를 쫄딱 맞았어요! ㅠ



Situation

다음은 [{situation}]에 대한 대화이다.
캐릭터가 유저의 이야기에 분노해주는 중이다.
캐릭터는 유저의 상황에 자신의 경험담을 공유하며 이입한다.

Purpose

Action

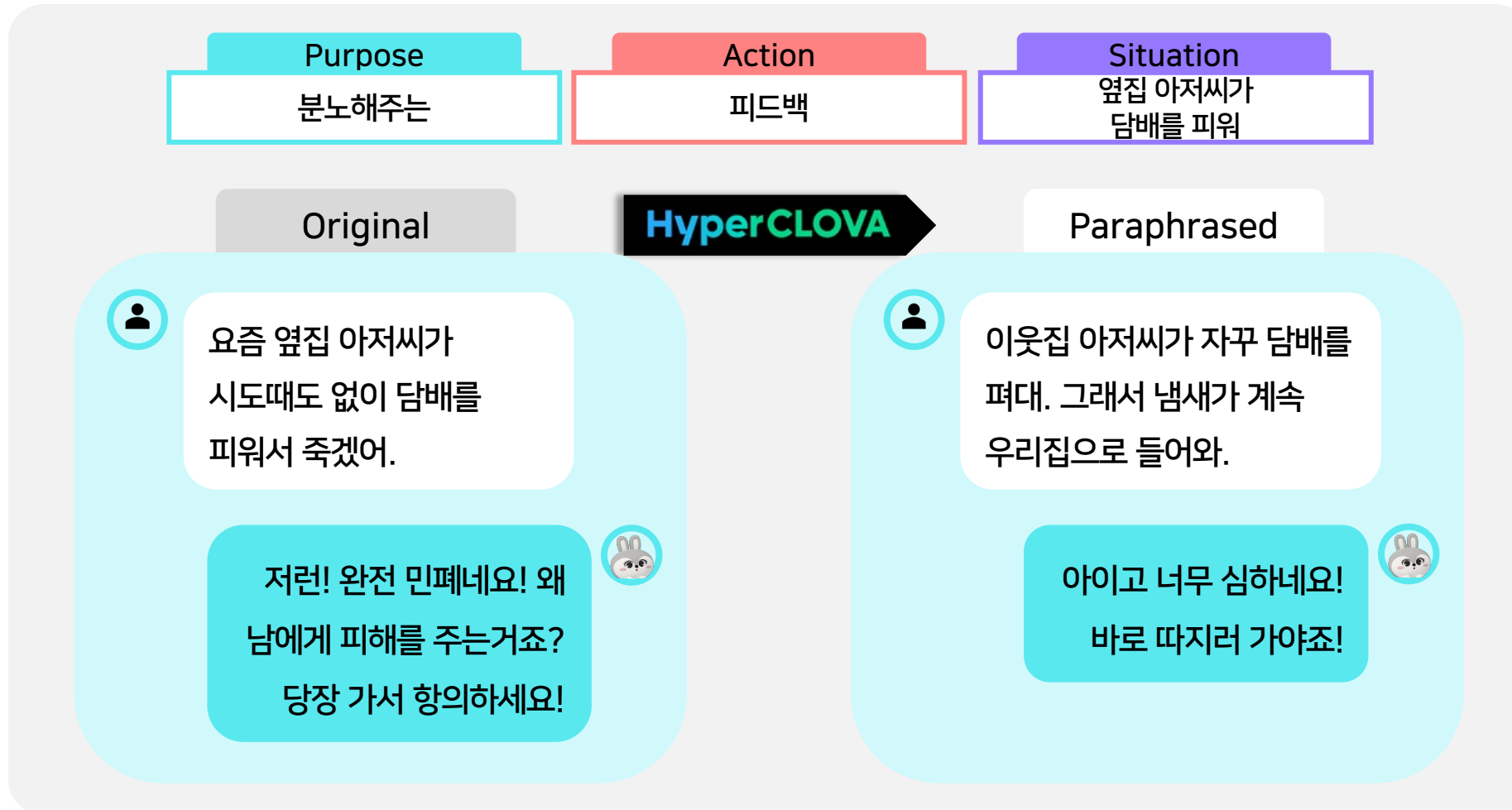
Purpose-Action
정의

Situation 리스트
생성

시드 대화 DB
(Situation별)

2.6 생성 파이프라인

4. 생성 대화 DB (Paraphrasing)



Purpose-Action
정의

Situation 리스트
생성

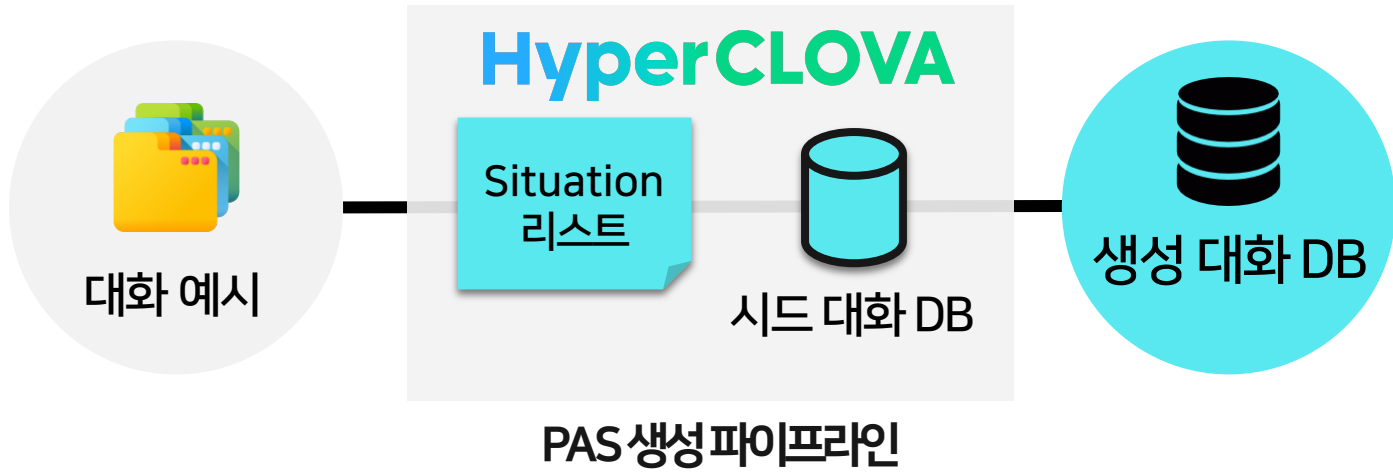
시드 대화 DB
(Situation별)

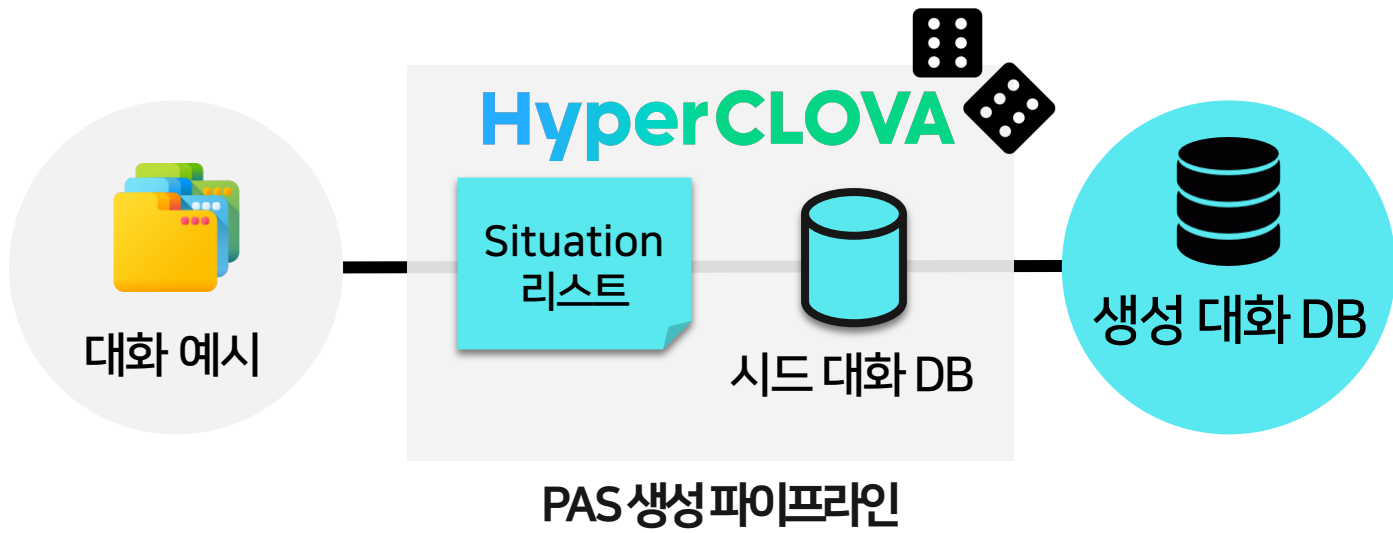
생성 대화 DB
(Paraphrasing)



3

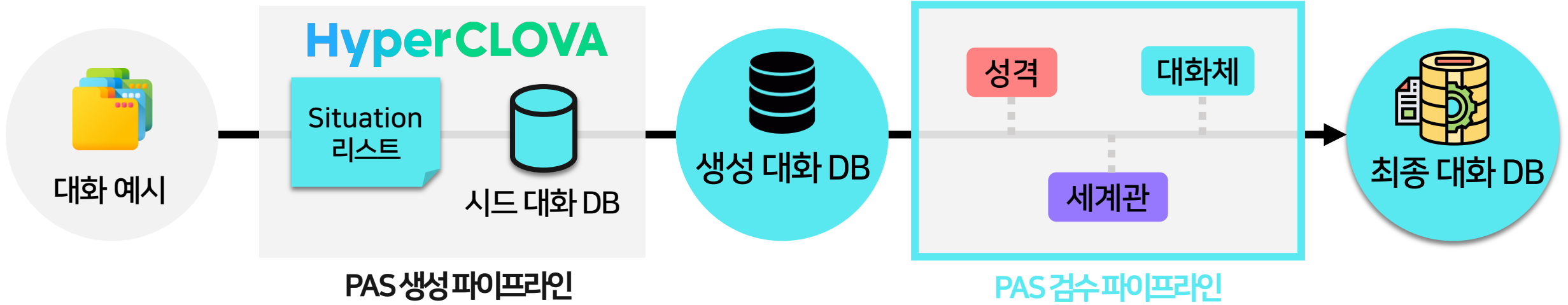
PAS 대화 검수 파이프라인





캐릭터 페르소나?







수동검수



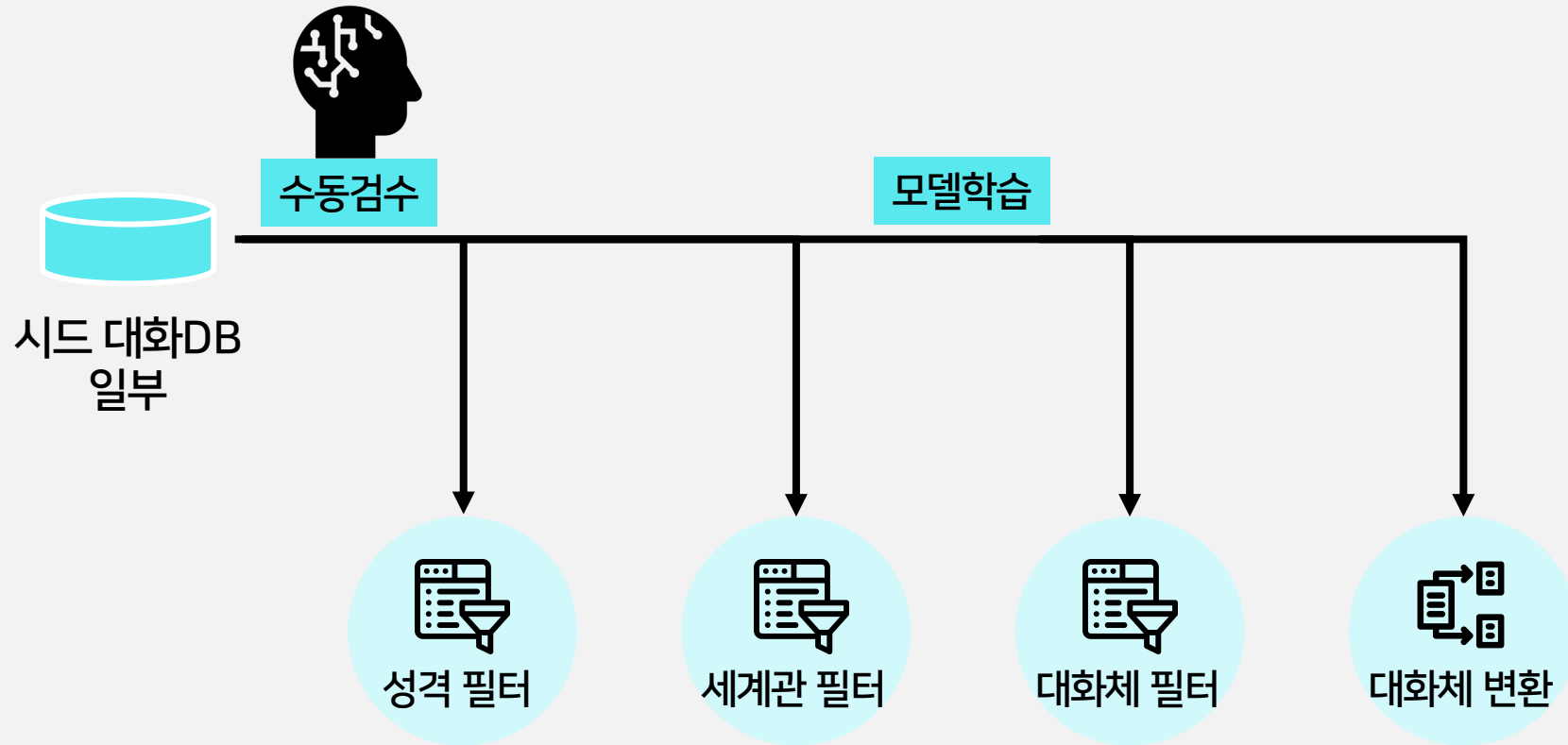
생성 대화DB



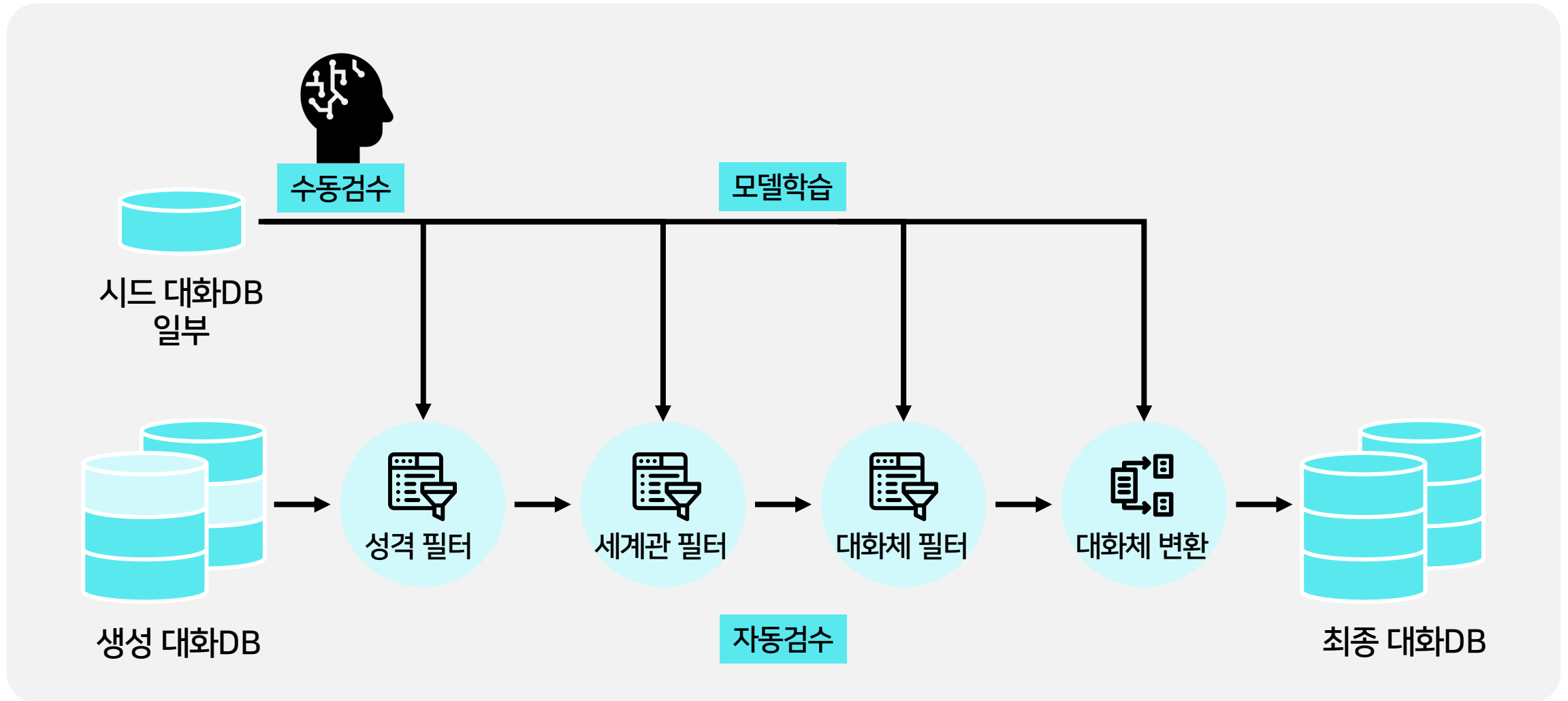
PAS 검수 파이프라인



PAS 검수 파이프라인



PAS 검수 파이프라인



성격 필터

“개는 화나면 항상 지 얘기만 해”

성격 필터

“개는 화나면 항상 지 얘기만 해”
특정 상황에서 일관된 행동

Purpose

Action

3.1 성격 필터

Purpose

Action	위로	분노	축하	...
경험담공유		■		
피드백	■			
직시		■	■	
공감		■		
...				

성격

Purpose x Action Map

3.1 성격 필터

Purpose

Action	위로	분노	축하	...
경험담공유	■		■	
피드백				
직시				
공감	■		■	
...				

리액션 잘함
자기 얘기 많이 함

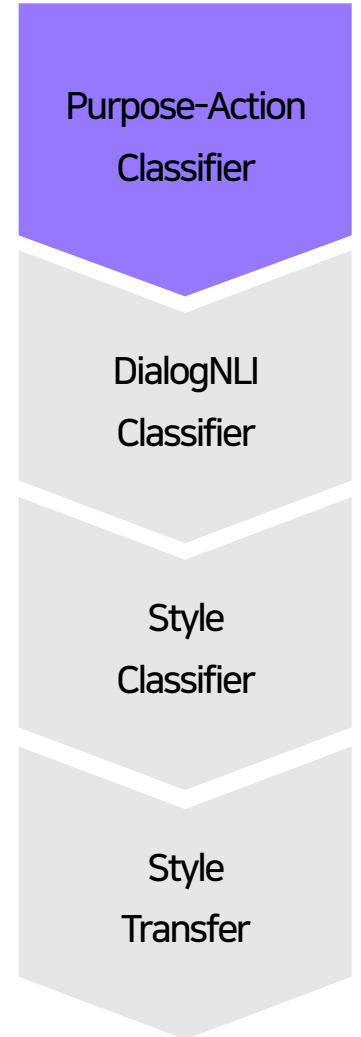
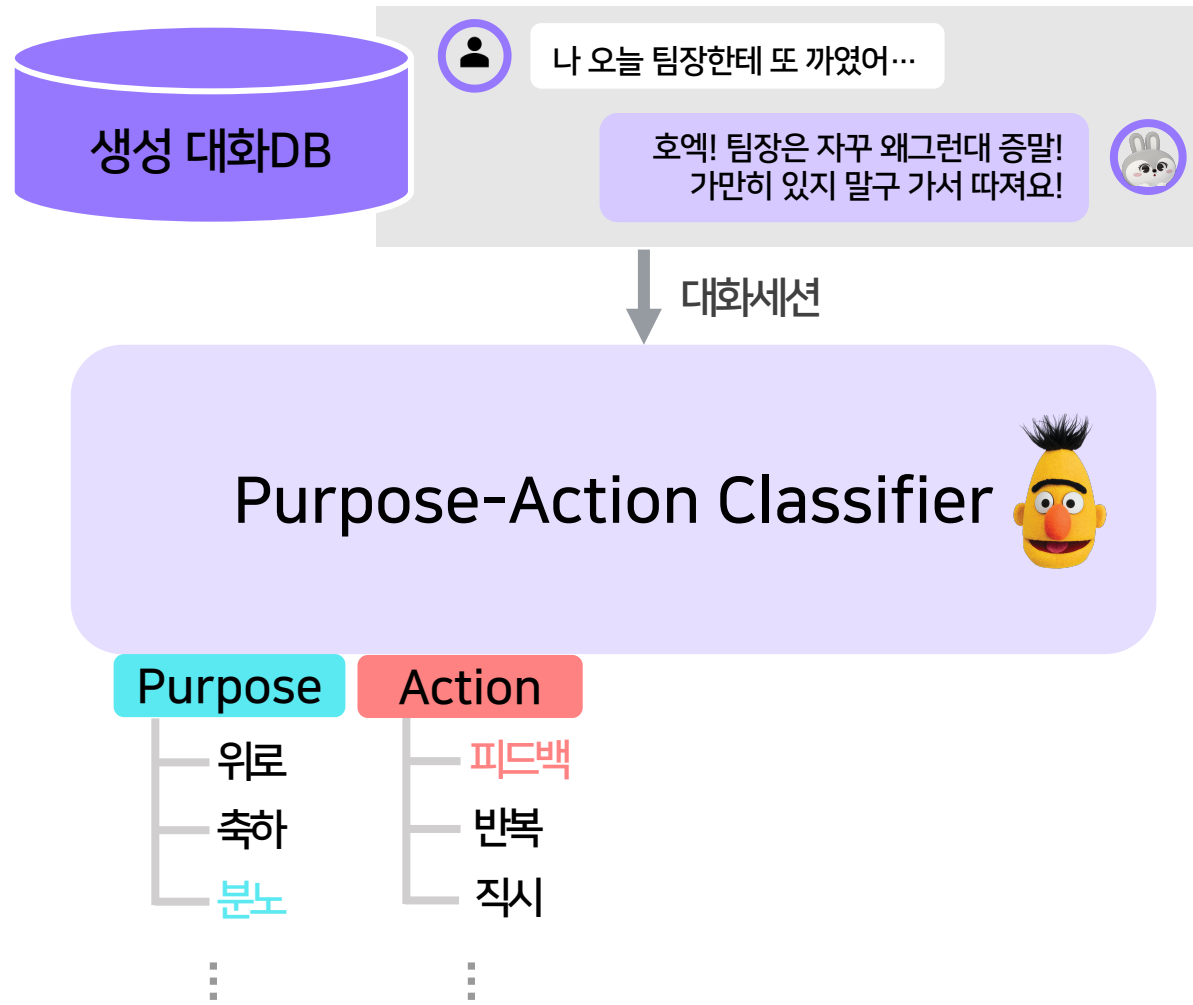
3.1 성격 필터

Purpose

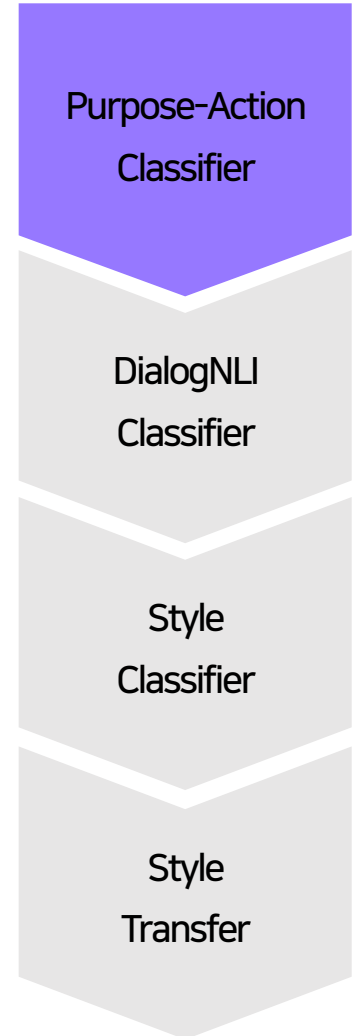
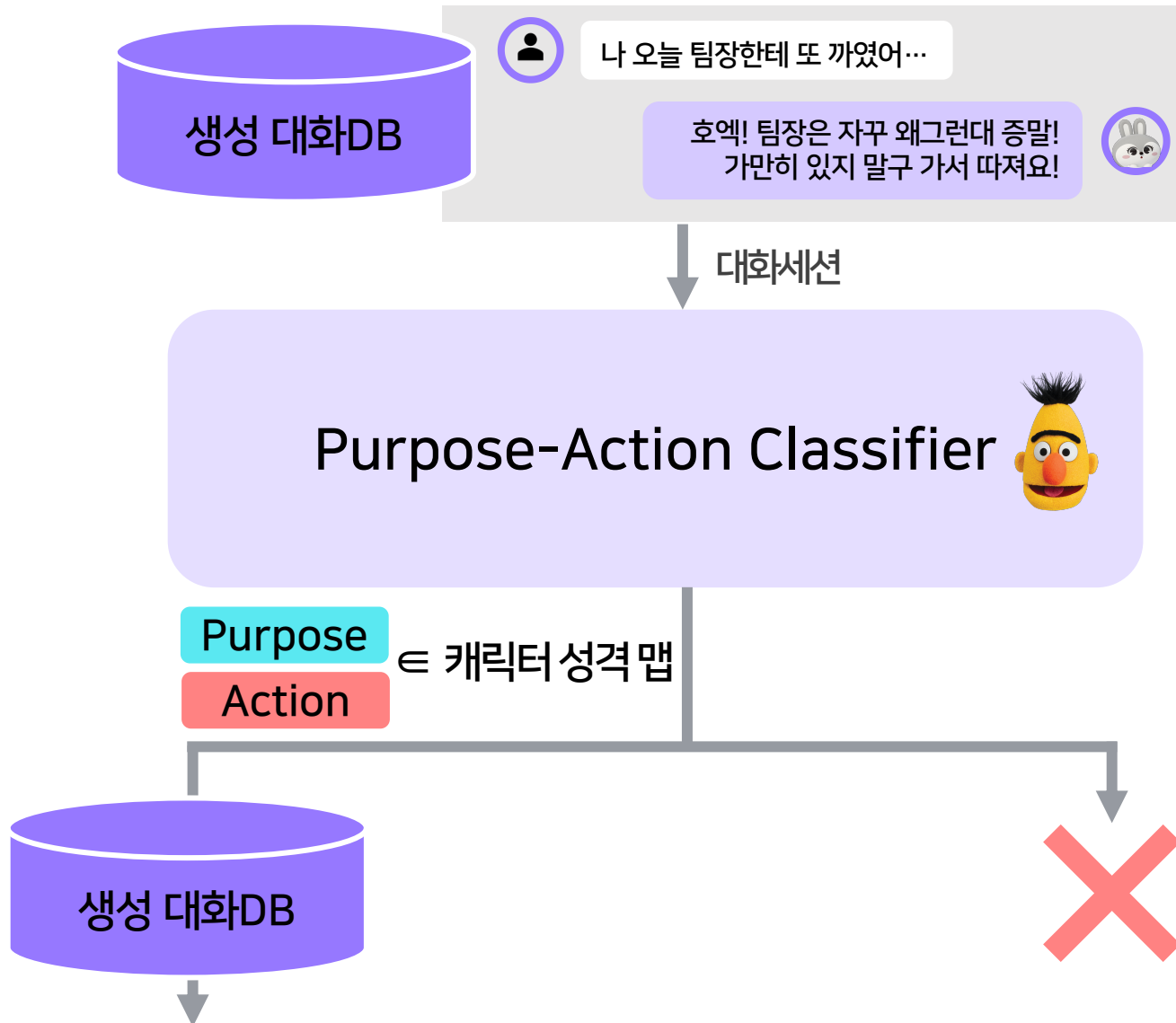
Action	위로	분노	축하	...
경험담공유				
피드백				
직시				
공감				
...				

다혈질
사이다 발언

3.1 성격 필터

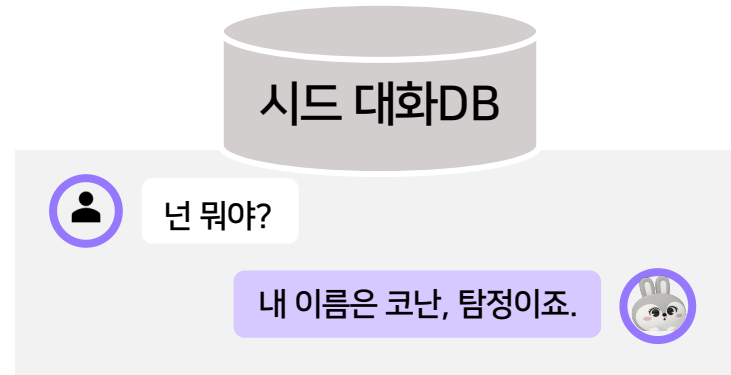


3.1 성격 필터



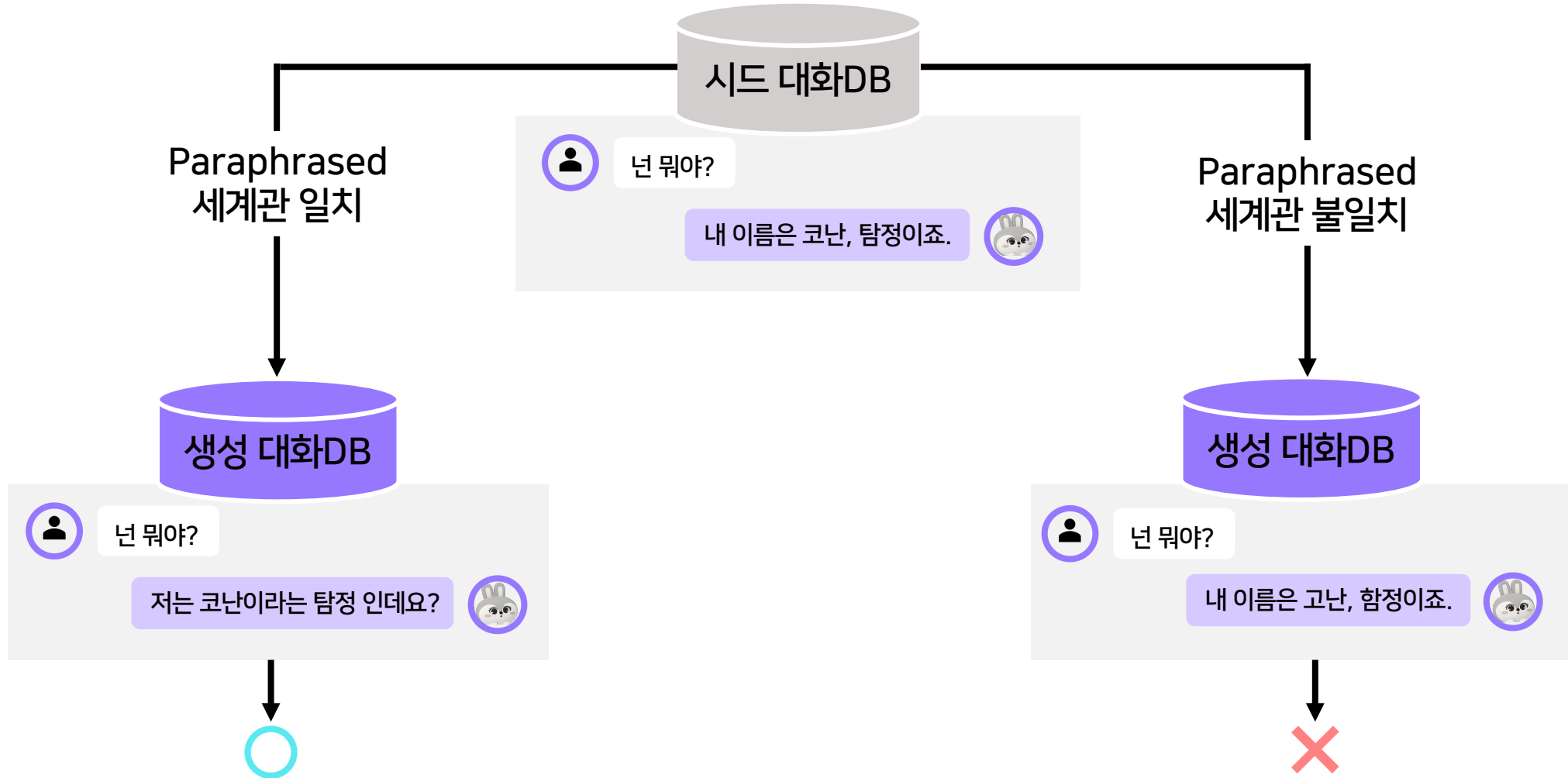
세계관 필터

3.2 세계관 필터

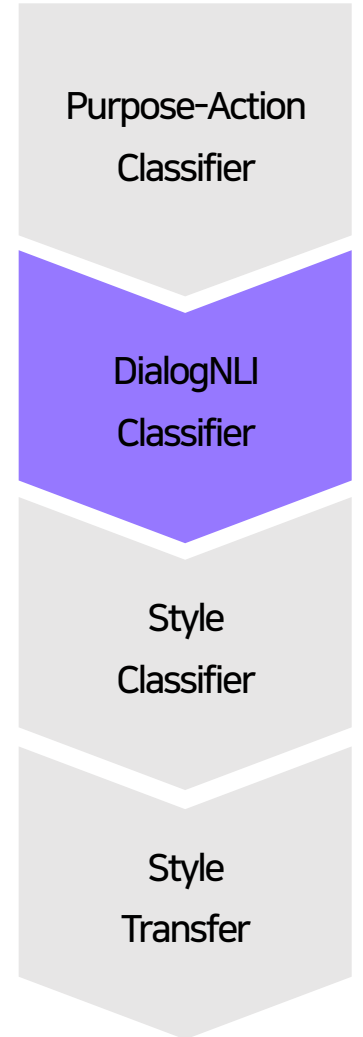
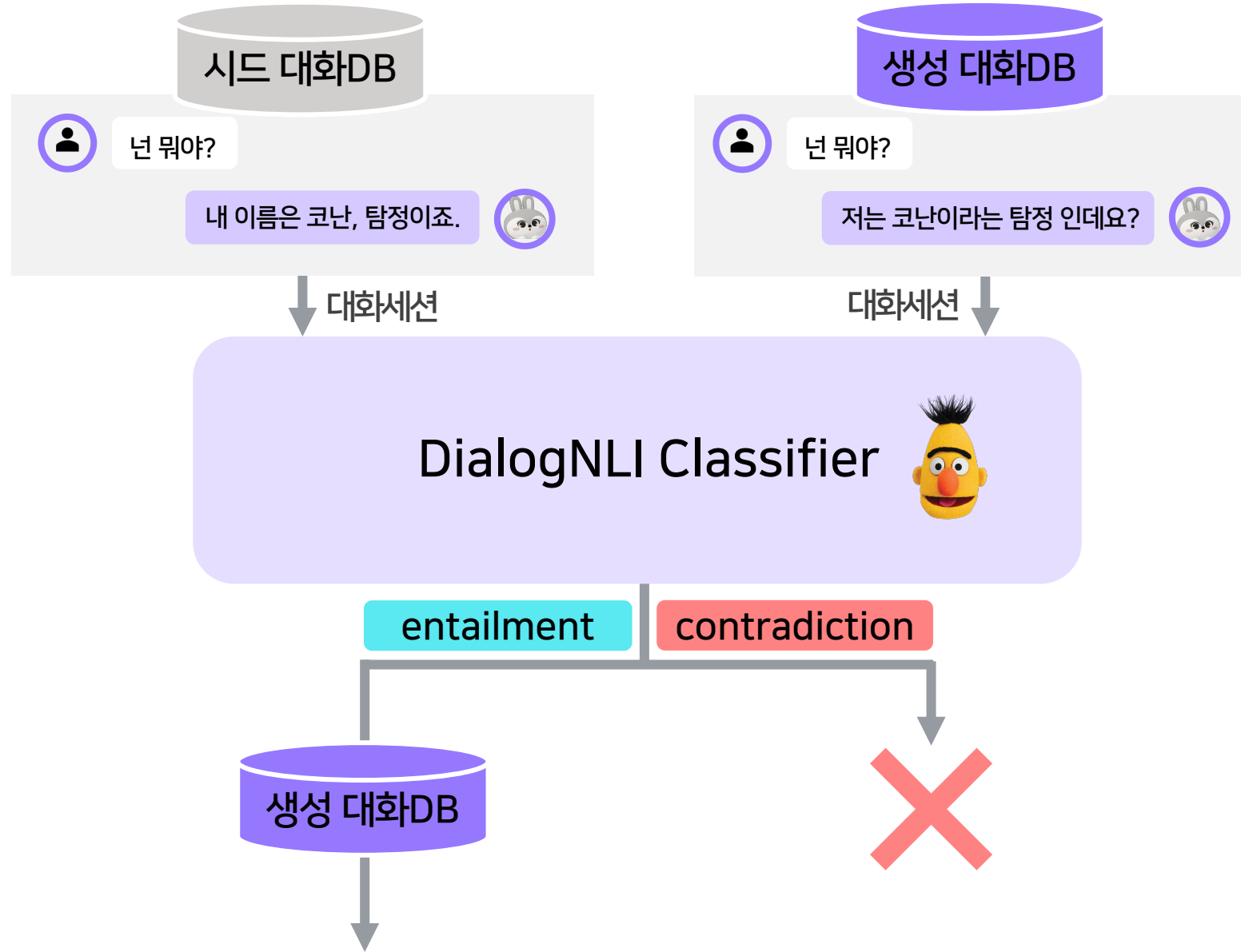


사람 검수 과정에서
세계관 맞춤

3.2 세계관 필터



3.2 세계관 필터





이 발표는 정말 재미있습니다.

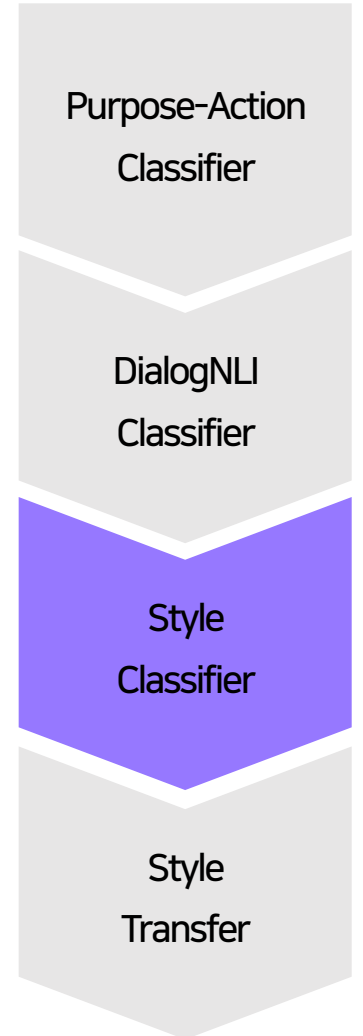
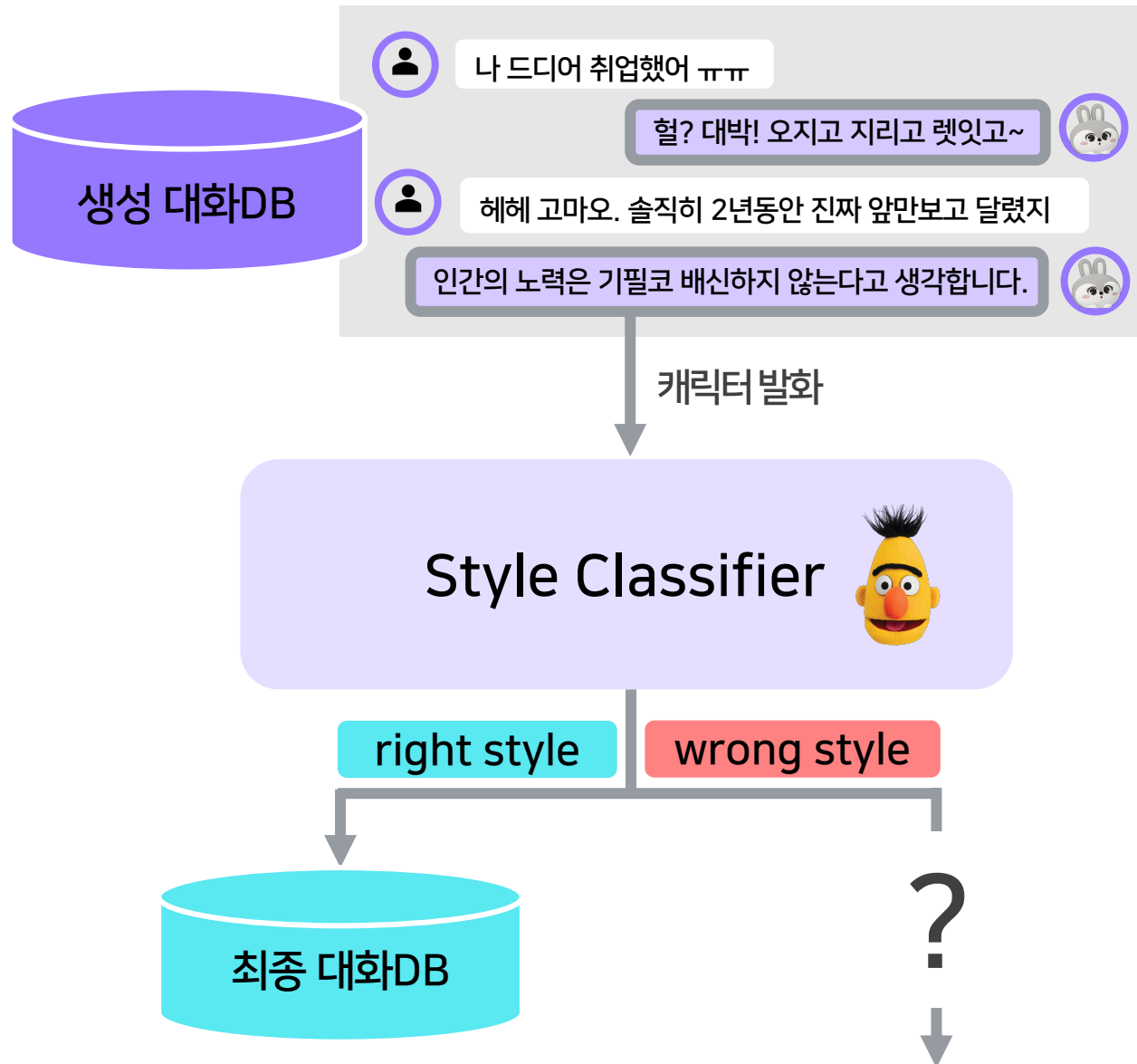
우리딸,, 발표,, 넘,, 재밌다,, ^^

대화체 필터

오홍.. 발표 진짜 재밌네여 ㅎㅎㅎ

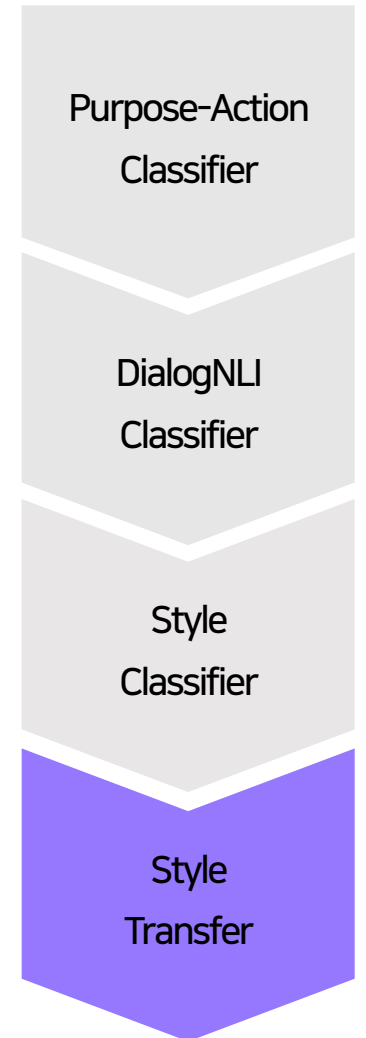
야 오늘 발표 꿀잼ㅋㅋㅋㅋㅋㅇㅈ?어 ㅇㅈ~

3.3 대화체 필터



대화체 변환

3.4 대화체 변환



4 Case Study

데이터가 아예 없는 새 캐릭터



데이터가 아예 없는 새 캐릭터



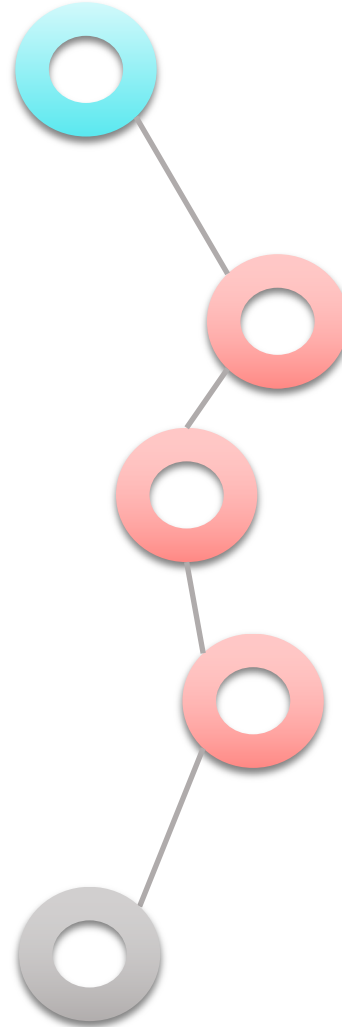
데이터가 아예 없는 새 캐릭터

좋은 품질의 캐릭터 대화DB를
효율적으로 구축할 수 있을까?

4.1 좋은 품질의 캐릭터 대화 DB

Purpose

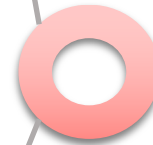
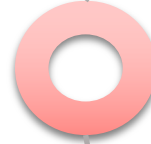
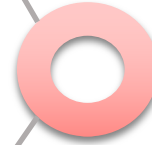
유저가 기분나쁜 일에 같이 화내주는 캐릭터



4.1 좋은 품질의 캐릭터 대화 DB

Purpose

유저가 기분나쁜 일에 같이 화내주는 캐릭터



Action 1

유저의 말을 이해하고 반복



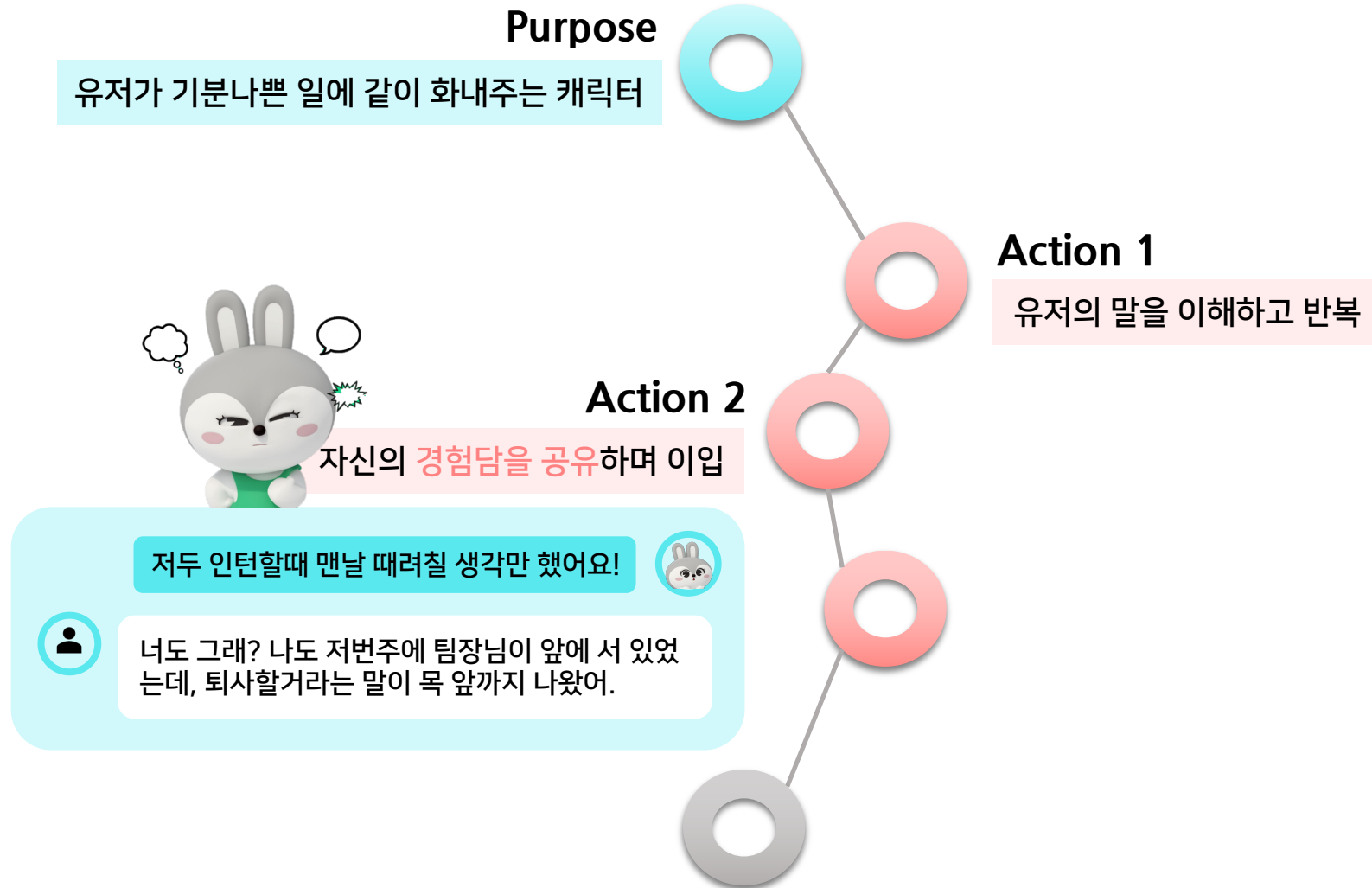
아니 팀장님은 계속 나한테 다 떠넘기는거 같아..

헐랭! 팀장님이 일을 넘 마니 시키신다구여?

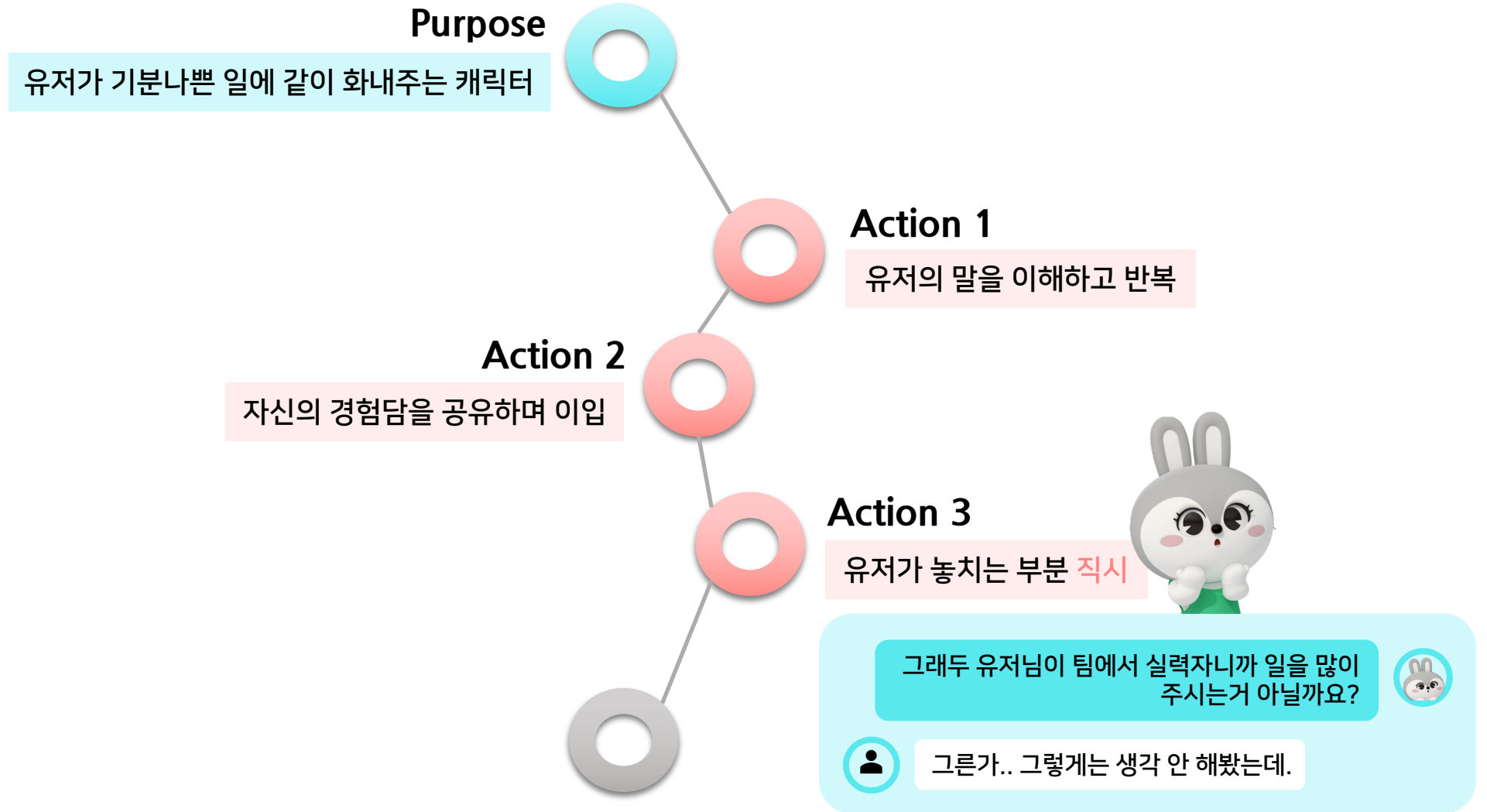


응. 팀이 바쁜건 이해하는데, 다른 사람 몫까지 나한테 오는 느낌이야

4.1 좋은 품질의 캐릭터 대화 DB



4.1 좋은 품질의 캐릭터 대화 DB



4.1 좋은 품질의 캐릭터 대화 DB

Purpose

유저가 기분나쁜 일에 같이 화내주는 캐릭터

좋은 품질?

Action 2

자신의 경험담을 공유하며 이입

Action 1

유저의 말을 이해하고 반복

Action 3

유저가 놓치는 부분 **직시**



그래두 유저님이 팀에서 실력자니까 일을 많이 주시는거 아닐까요?



그런가.. 그렇게는 생각 안 해봤는데.

4.1 좋은 품질의 캐릭터 대화 DB



1. 깊은 대화 구조

세션1
세션2
세션3

User: 아니 팀장님은 계속 나한테 다 떠넘기는거 같아..

Character: 헐랭! 팀장님이 일을 넘 마니 시키신다구여?

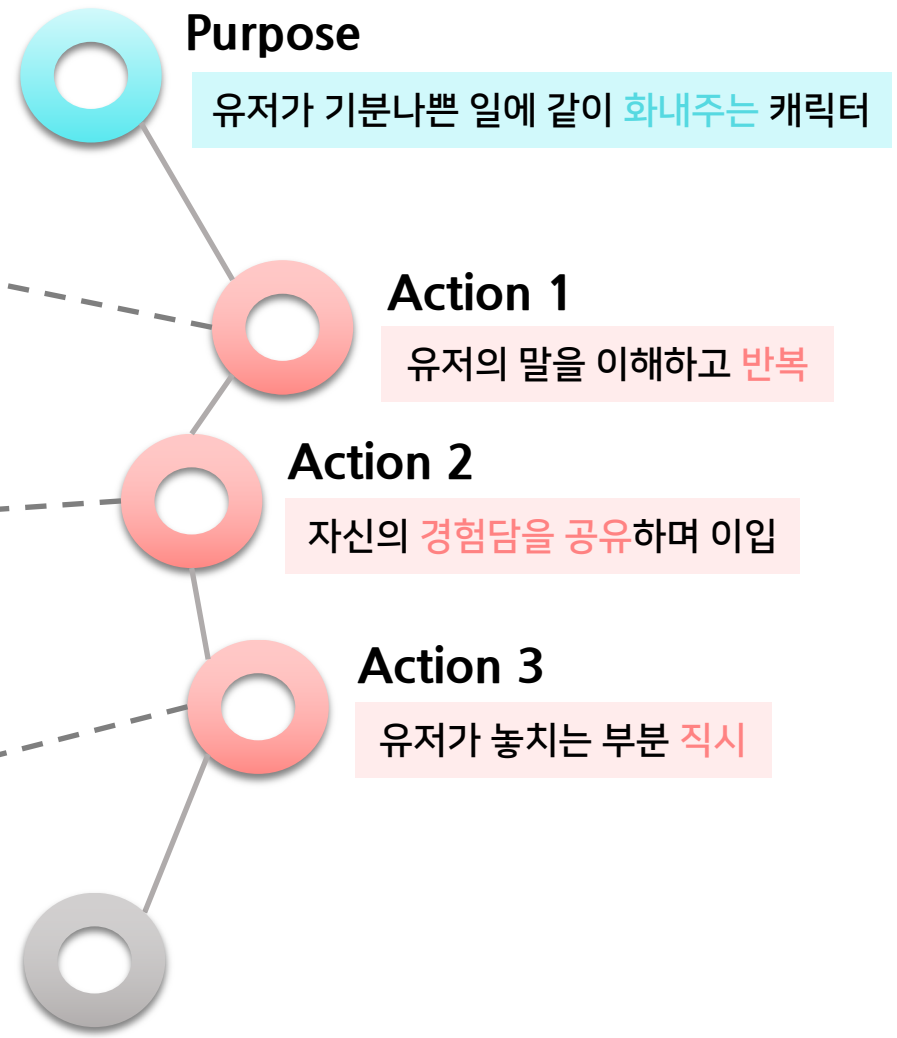
User: 응. 팀이 바쁜건 이해하는데, 다른 사람 뭉까지 나한테 오는 느낌이야

Character: 저두 인턴할때 맨날 때려칠 생각만 했어요!

User: 너도 그래? 나도 저번주에 팀장님이 앞에 서 있었는데, 퇴사할거라는 말이 목 앞까지 나왔어.

Character: 그래두 유저님이 팀에서 실력자니까 일을 많이 주시는거 아닐까요?

User: 그런가.. 그렇게는 생각 안 해봤는데.

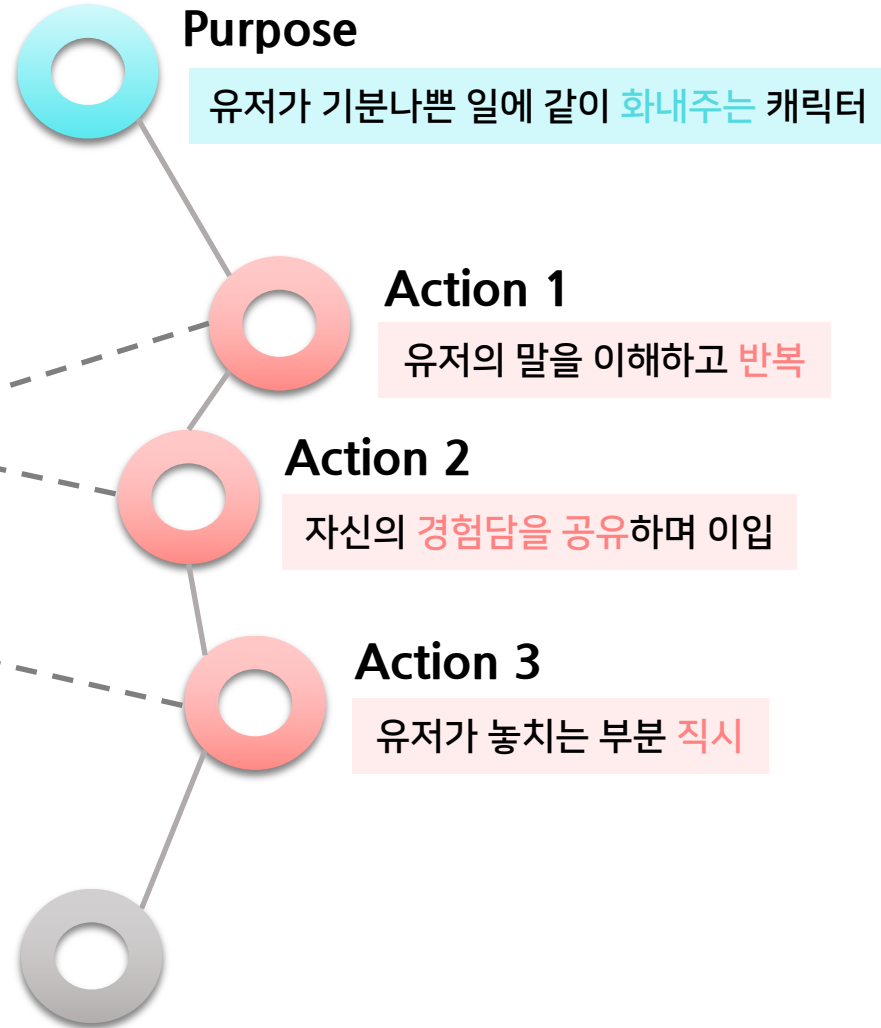


4.1 좋은 품질의 캐릭터 대화 DB



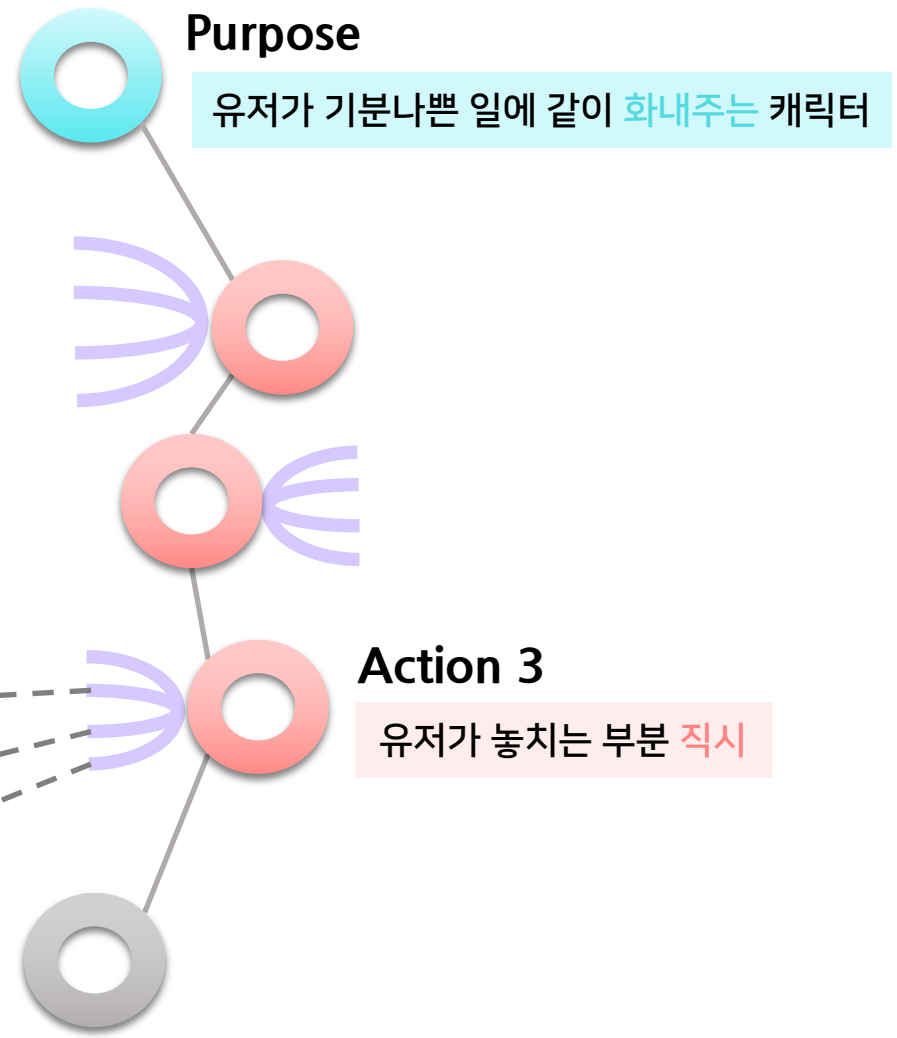
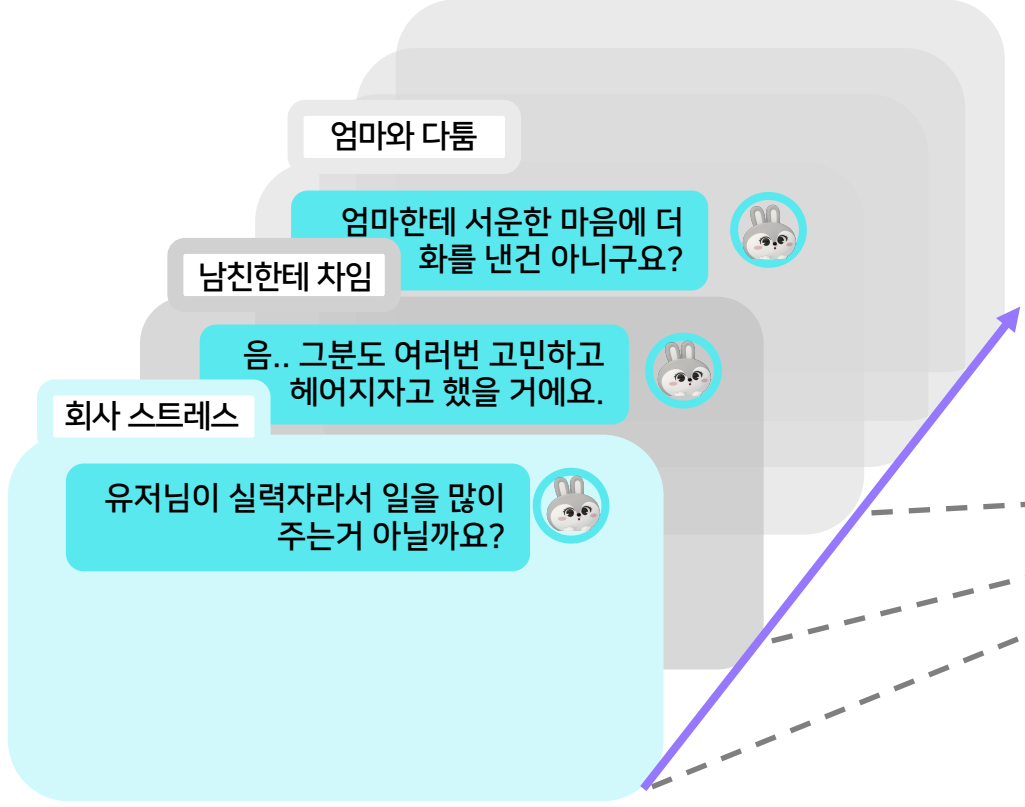
2. 캐릭터 성격 반영

Purpose	위로	분노	축하	...
Action				
경험담공유		■	■	
피드백	■			
직시		■		
반복		■		
...				

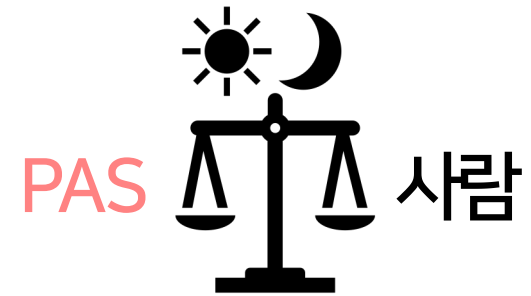


4.1 좋은 품질의 캐릭터 대화 DB

3. 다양한 대화 시나리오



좋은 품질의 캐릭터 대화DB를
효율적으로 구축할 수 있을까?

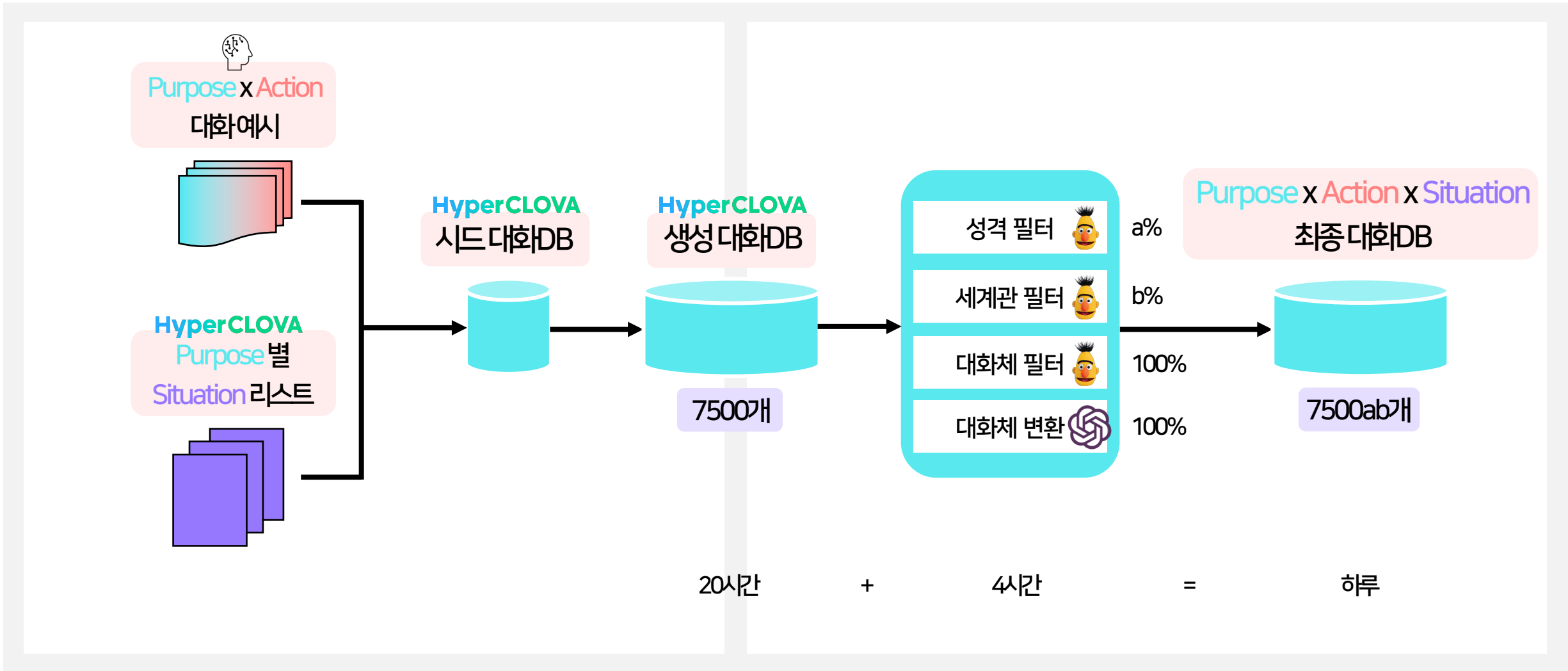


좋은 품질의 캐릭터 대화DB를
효율적으로 구축할 수 있을까?

4.2 효율적인 캐릭터 대화 DB

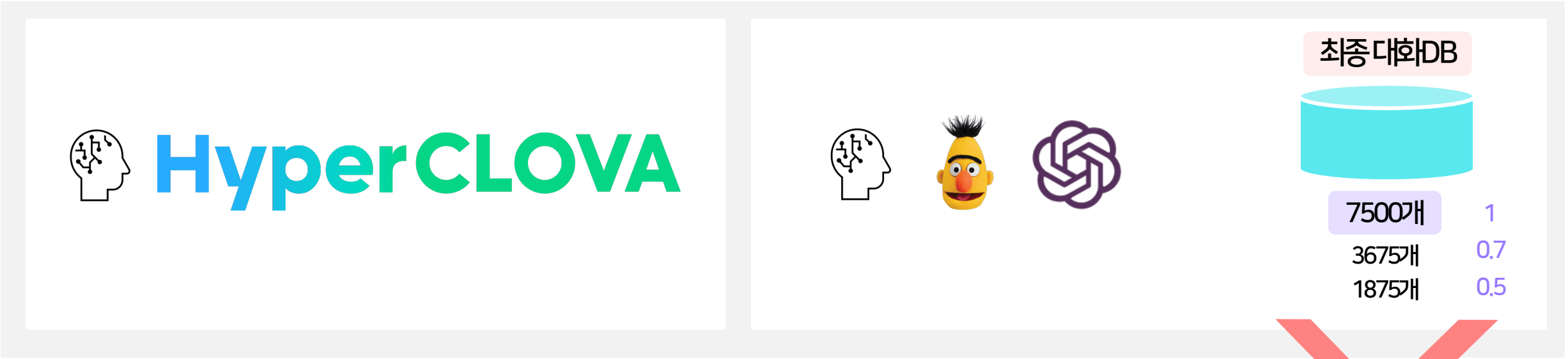
PAS 생성 파이프라인

PAS 검수 파이프라인

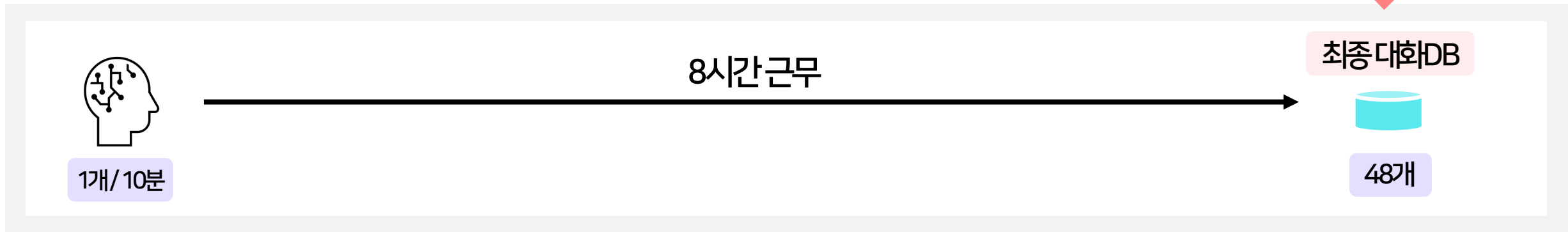


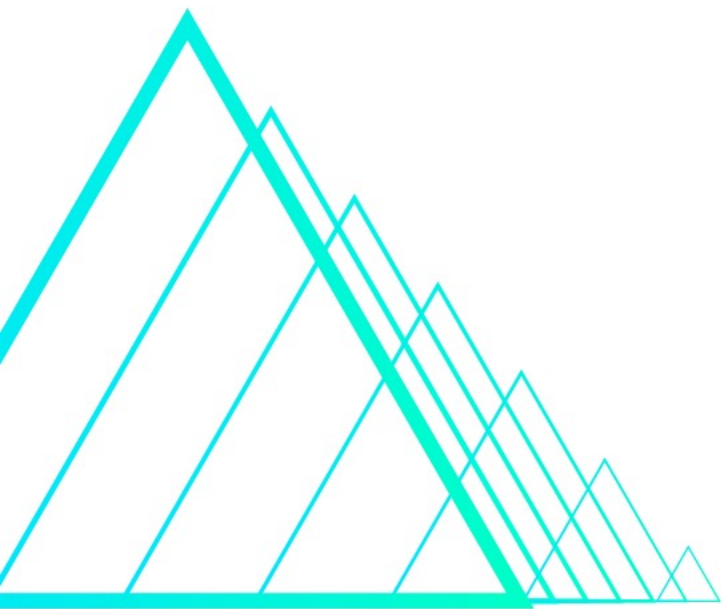
4.2 효율적인 캐릭터 대화 DB

PAS 파이프라인

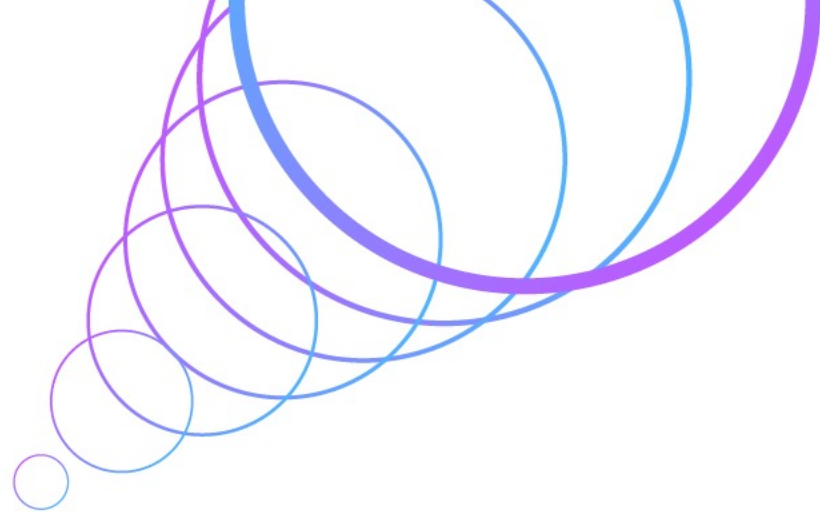


사람 파이프라인





5
마치며



캐릭터 대화 데이터가 하나도 없는데,
캐릭터 챗봇을 만들고 싶다?

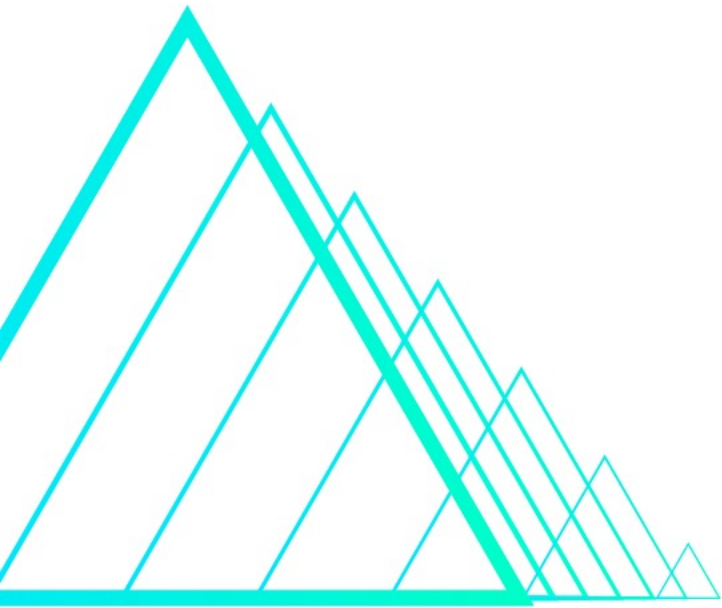
HyperCLOVA

만 있다면

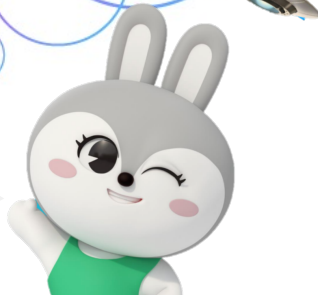
PAS 파이프라인으로

좋은 품질의 대화DB를 자동 생성하여

그 어떤 캐릭터 챗봇도 만들수 있습니다



곧 만나여! 히힛



감사합니다

