

#자연어

#대화  
시스템

# 현직자가 주목하고 있는 대화 시스템 기술 및 인사이트

김원우 강사



# CONTENTS

## CHAPTER 1

### 자연어 이해

#### 01. NLU 성능 향상

- 1) Joint Training
- 2) SLU

## CHAPTER 2

### 대화 시스템

#### 01. DST

- 1) Unseen Slot

#### 02. DM

- 1) Query Reformulation

## CHAPTER 3

### 인사이트

#### 01. GPT 3 와 대화 시스템

#### 02. Blender bot 2.0

CHAPTER 1

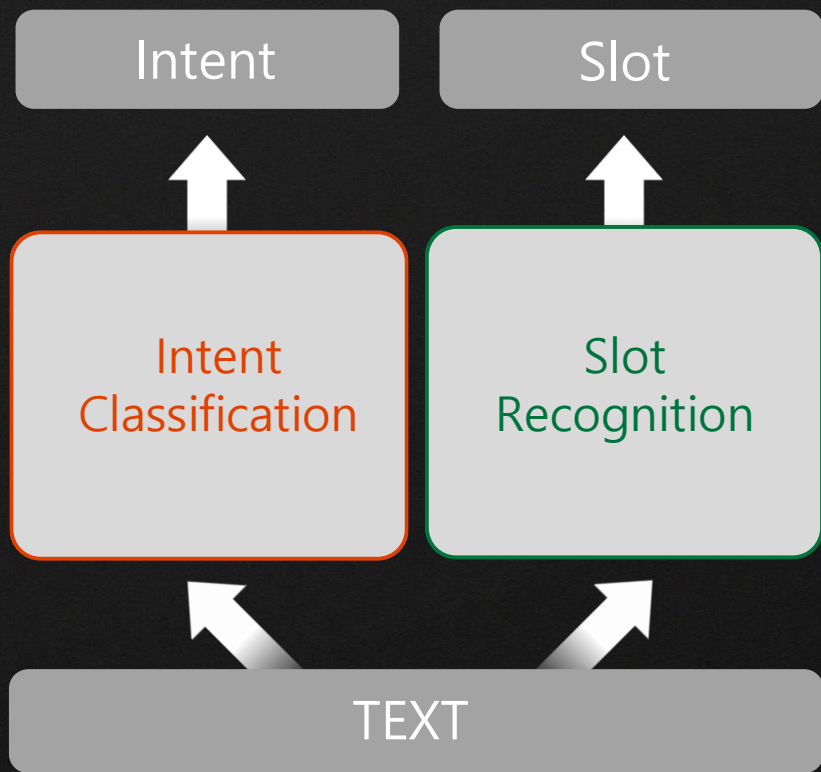
# 자연어 이해



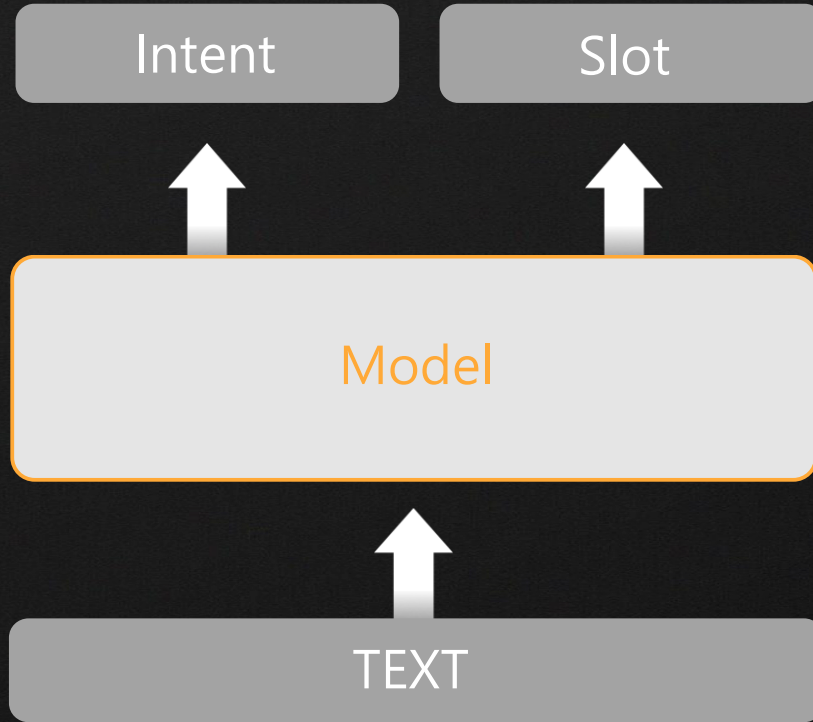
1

# Joint Training

## 기존의 NLU

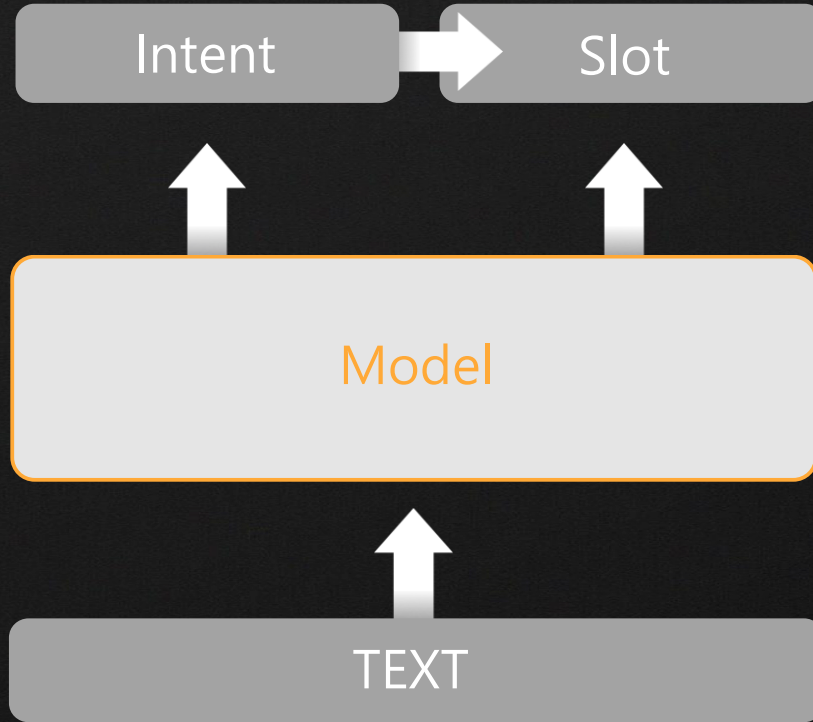


## Joint Training

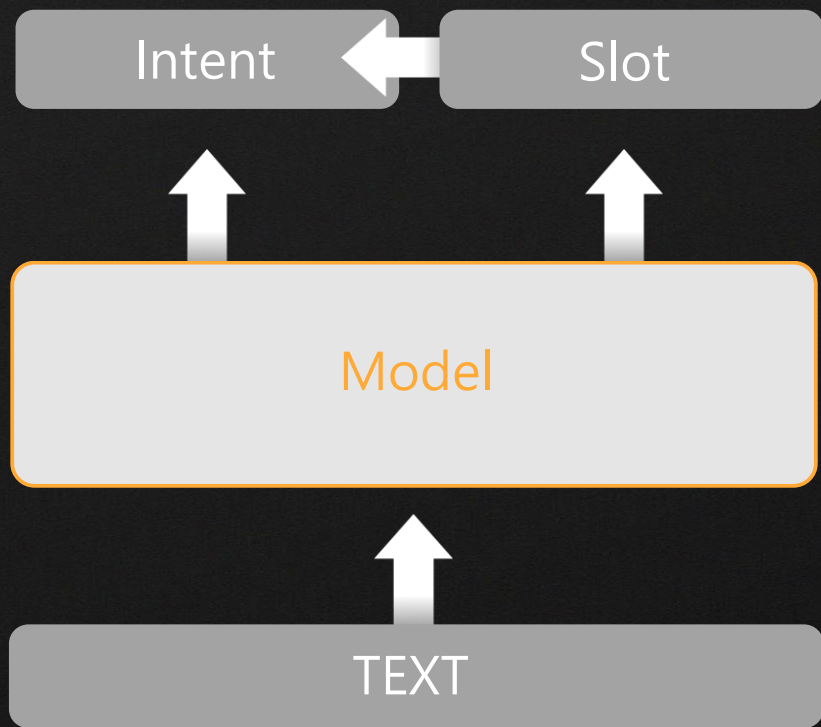




## Joint Training

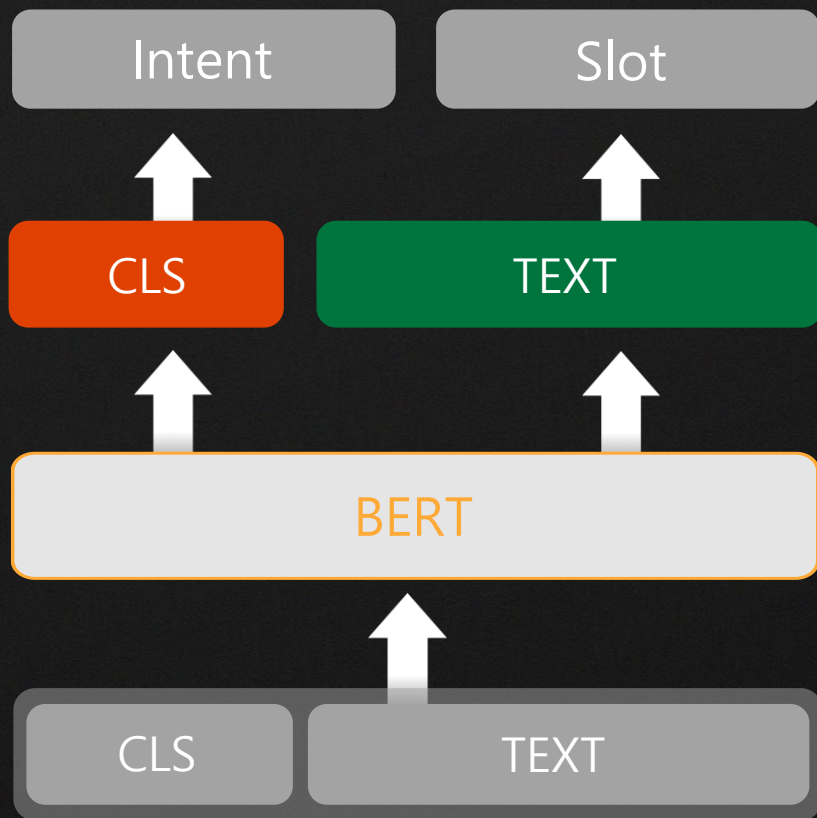


## Joint Training

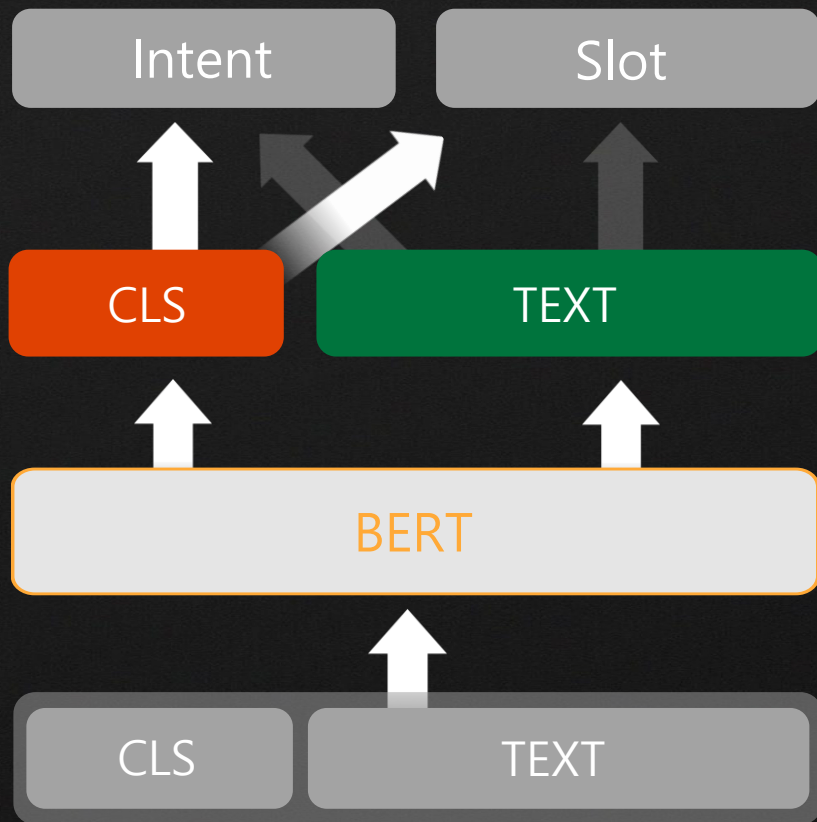




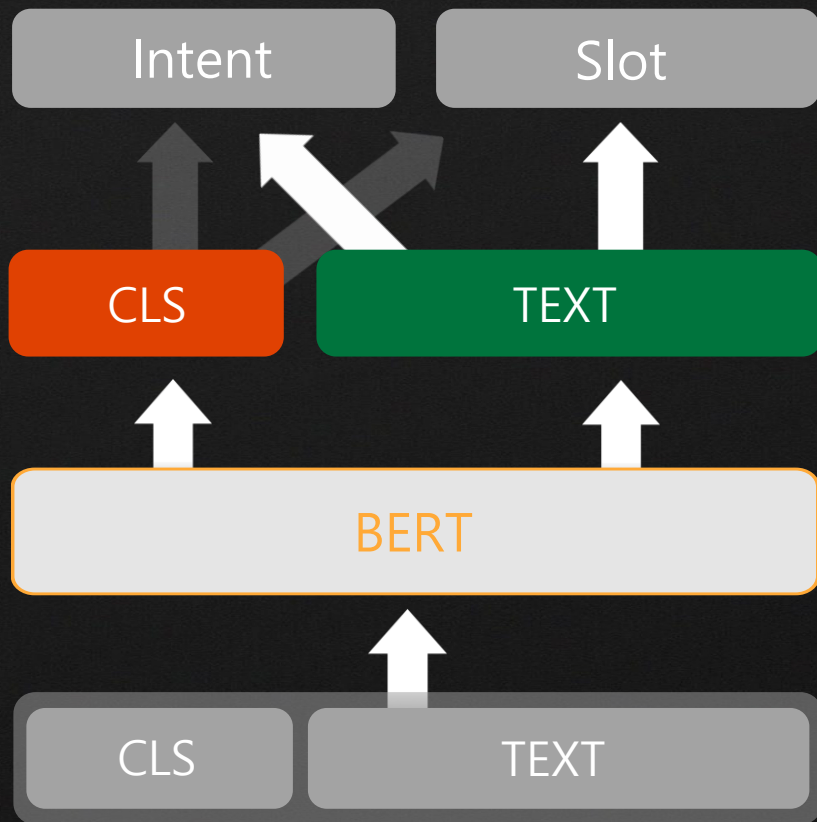
## Joint Training with BERT : BERT for Joint Intent Classification and Slot Filling



Cross Gated Joint Training : 자연어 이해 모델의 성능 향상을 위한 교차 게이트 메커니즘 방



Cross Gated Joint Training : 자연어 이해 모델의 성능 향상을 위한 교차 게이트 메커니즘 방

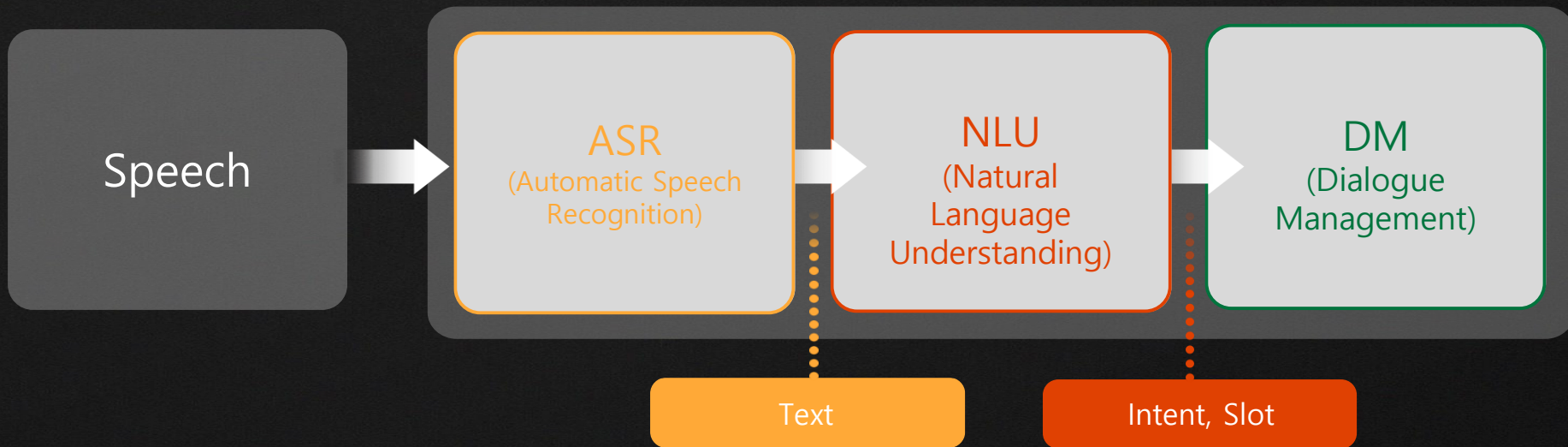




2

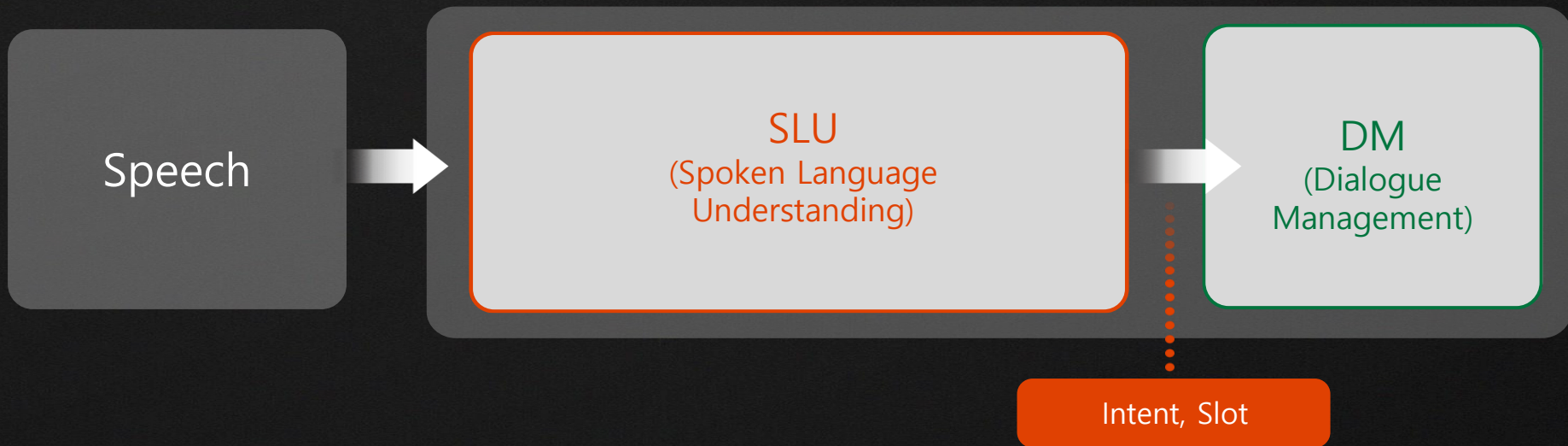
# Spoken Language Understanding

# Spoken Language Understanding



# Spoken Language Understanding :

## Speech Model Pre-training for End-to-End Spoken Language Understanding





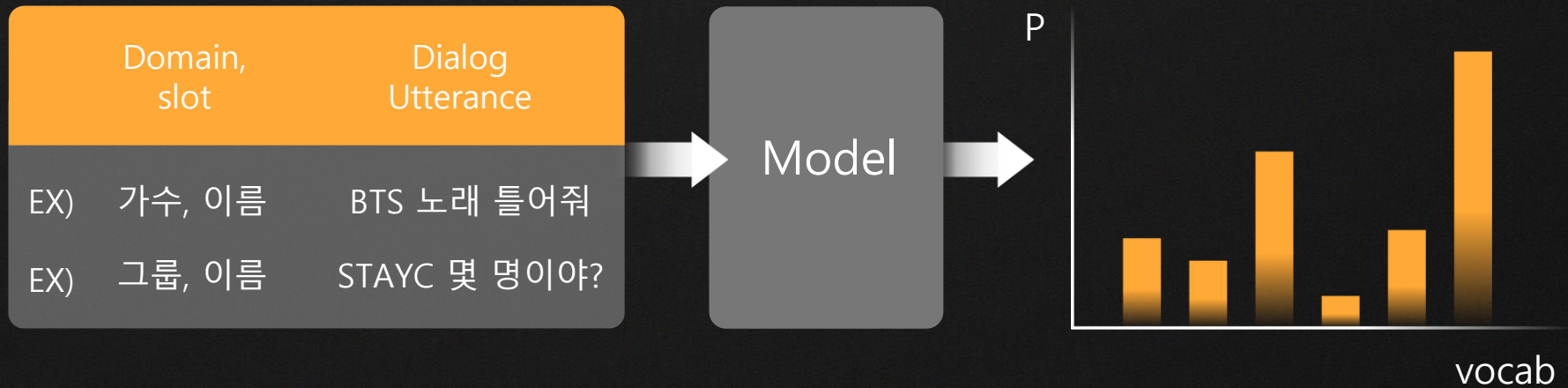
## CHAPTER 2

# 대화 시스템

1

# Dialog State Tracking

## 기존의 E2E DST





## Unseen Slot : 특정 도메인 대화가 가능한 봇에게 다른 도메인 질문할 경우 생기는 문제



BTS 노래 틀어줘

BTS 노래 BUTTER를 틀어 드릴게요.



Domain : 가수  
Slot : 이름



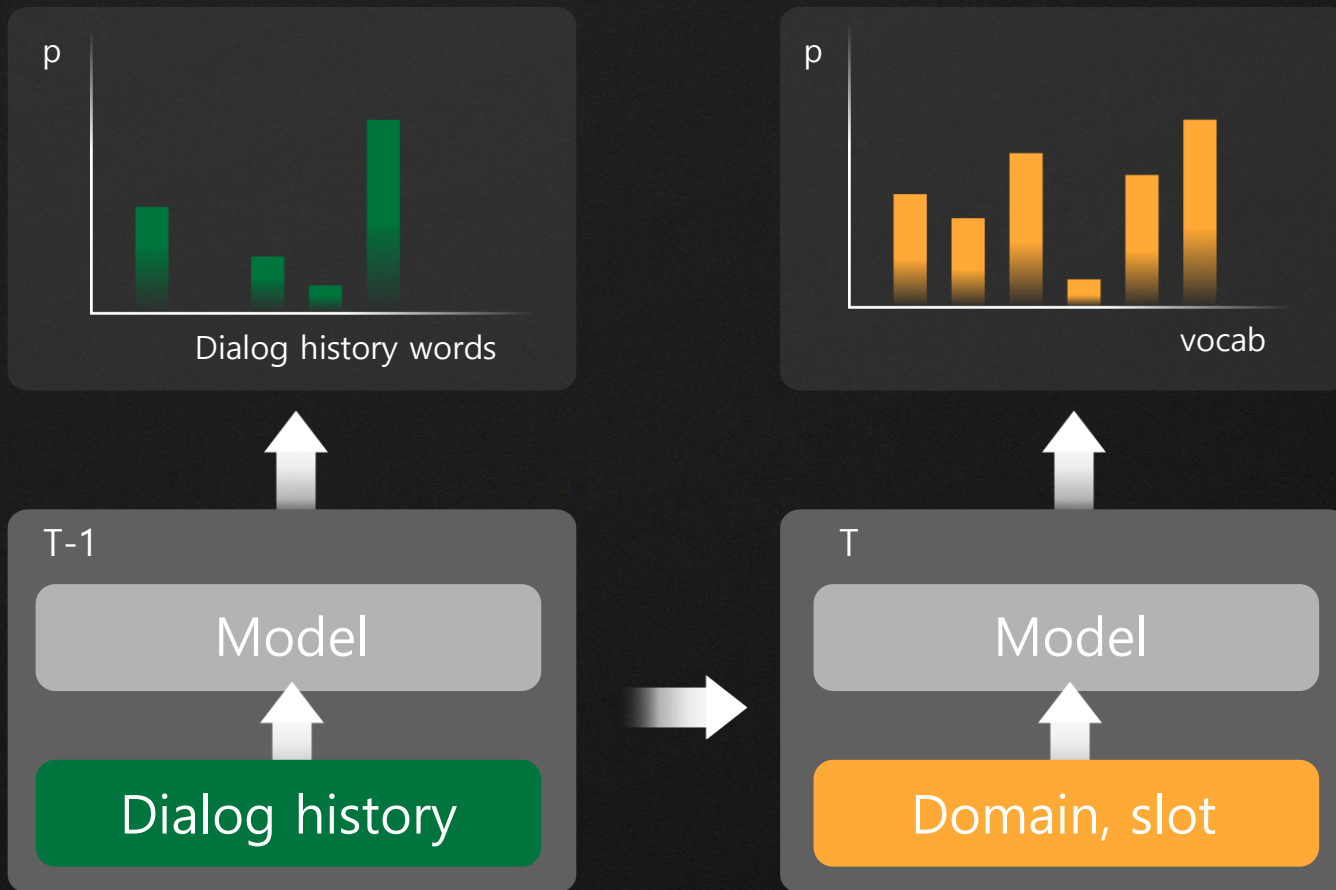
STAYC 몇 명이야?

(그룹 정보를 물어보는 학습은 너무 적게 해서  
STAYC 는 무슨 뜻일까...)

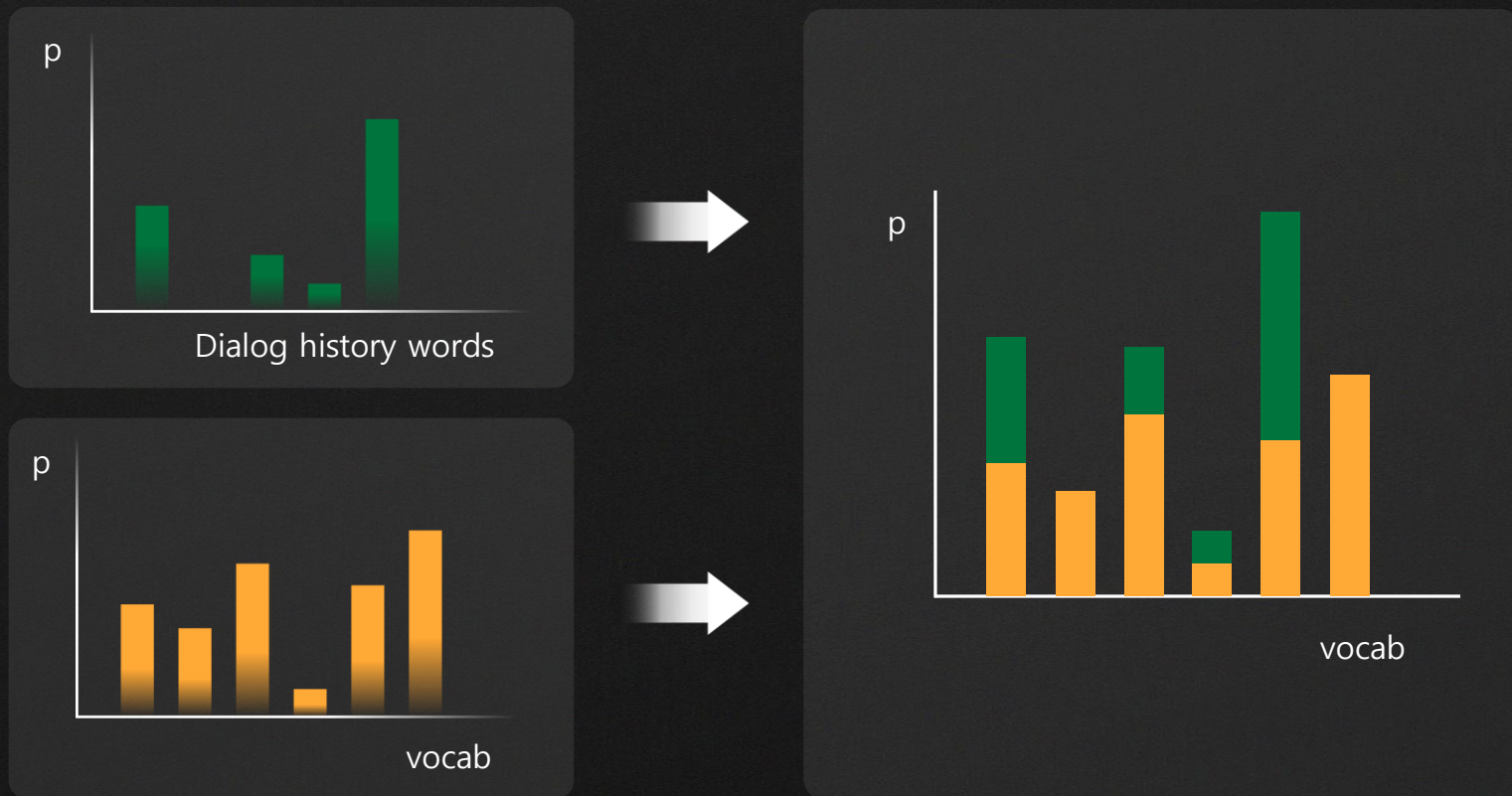


Domain : 그룹  
Slot : 이름

# Unseen Slot : Transferable Multi-Domain State Generator for Task-Oriented Dialogue Systems



# Unseen Slot : Transferable Multi-Domain State Generator for Task-Oriented Dialogue Systems

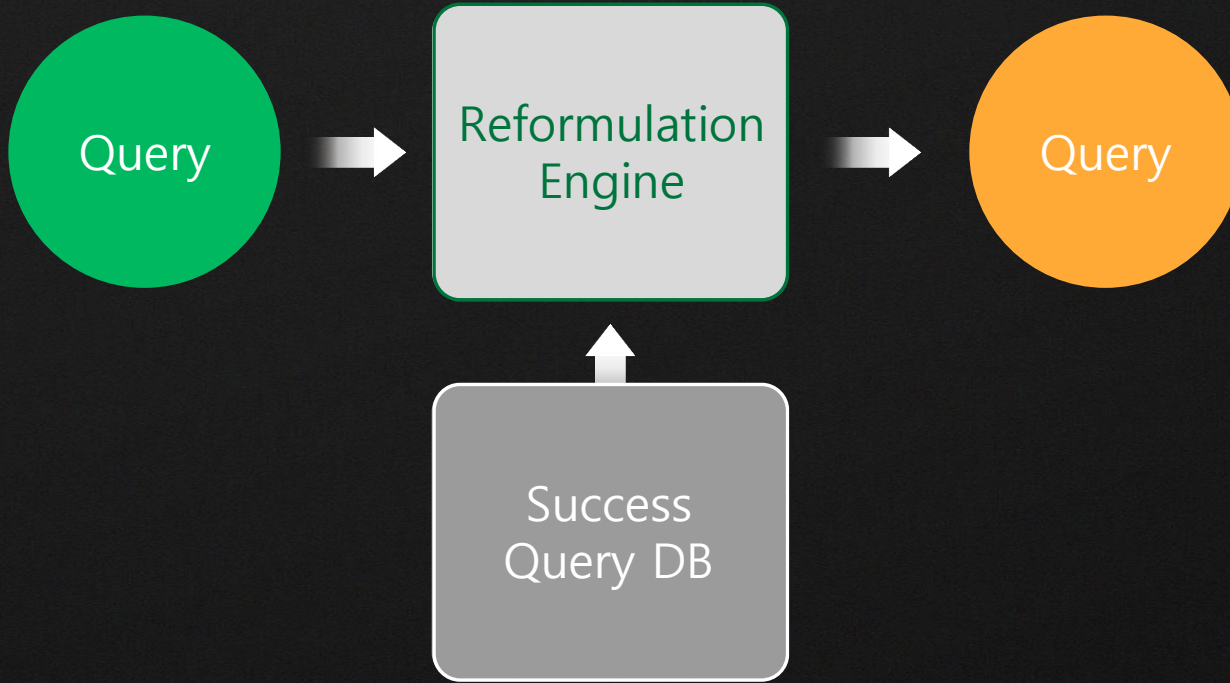




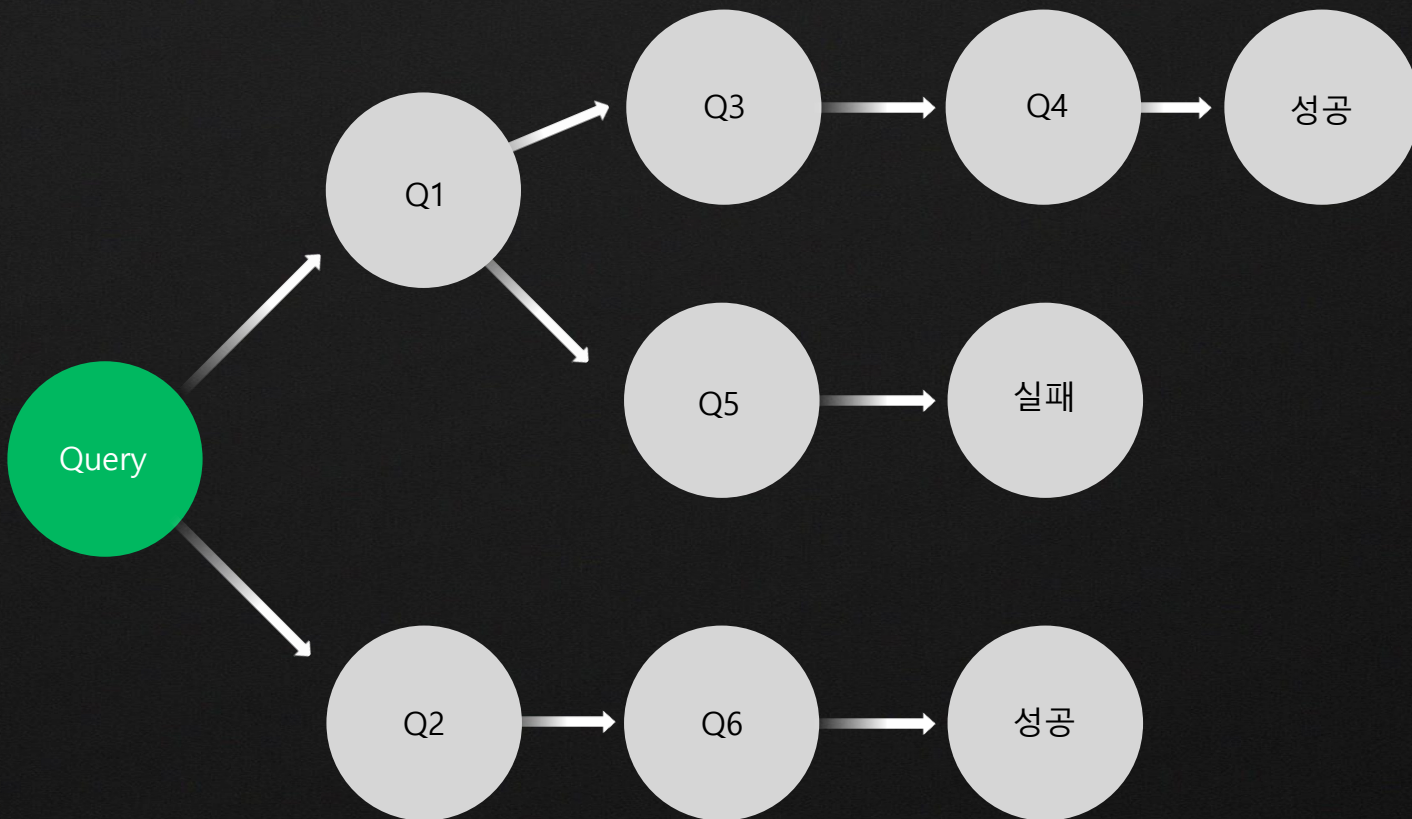
2

# Query Reformulation

## Query reformulation



# Query reformulation : Feedback-Based Self-Learning in Large-Scale Conversational AI Agents





## Query reformulation : Feedback-Based Self-Learning in Large-Scale Conversational AI Agents

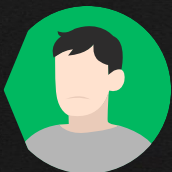


강산에 '거꾸로 강을 거슬러' 틀어줘

무슨 말인지 모르겠어요.



실패



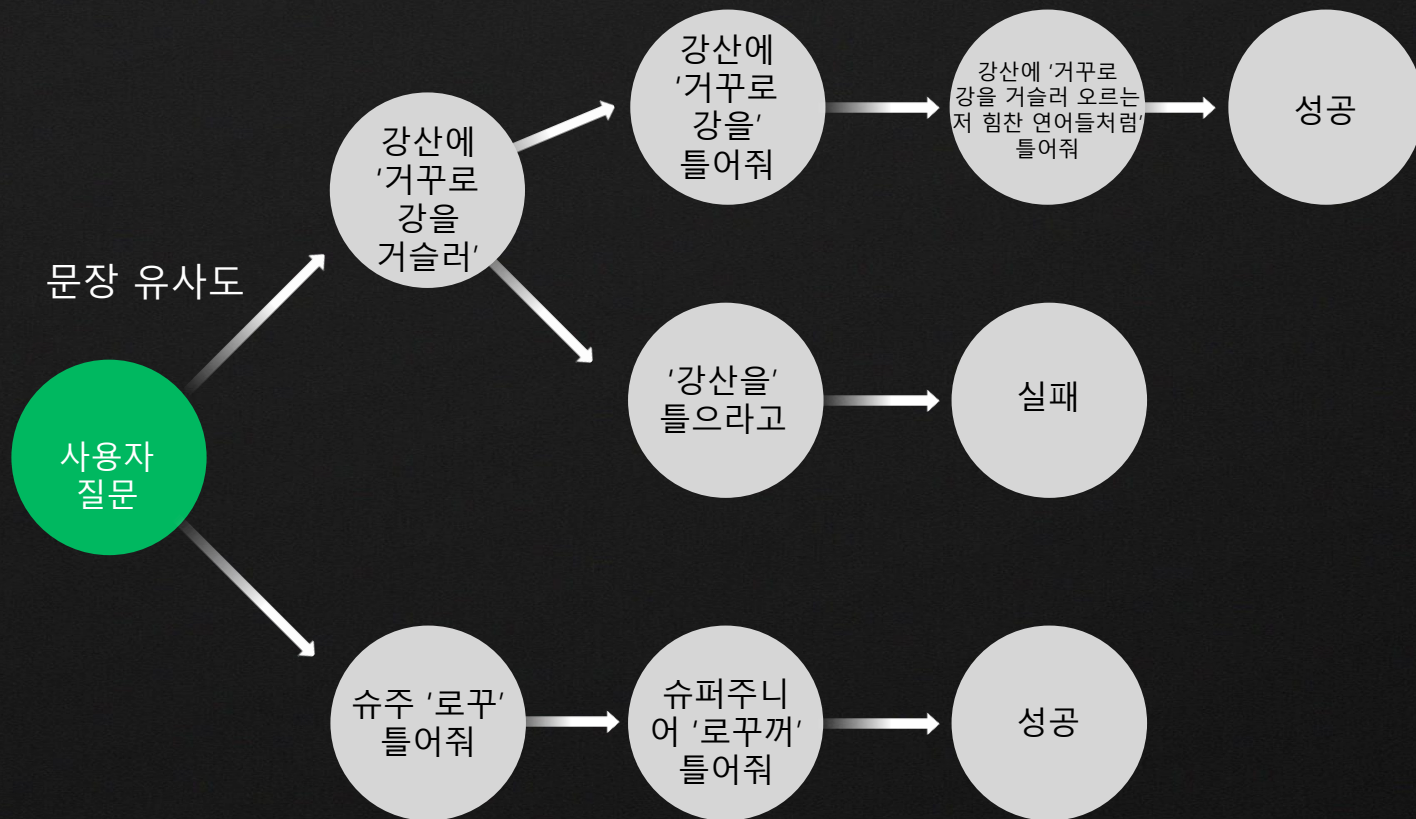
강산에 '거꾸로 강을 거슬러  
오르는 저 힘찬 연어들처럼'  
틀어줘

강산에의 '거꾸로 강을 거슬러 오르는  
저 힘찬 연어들처럼' 재생 해드릴게요.

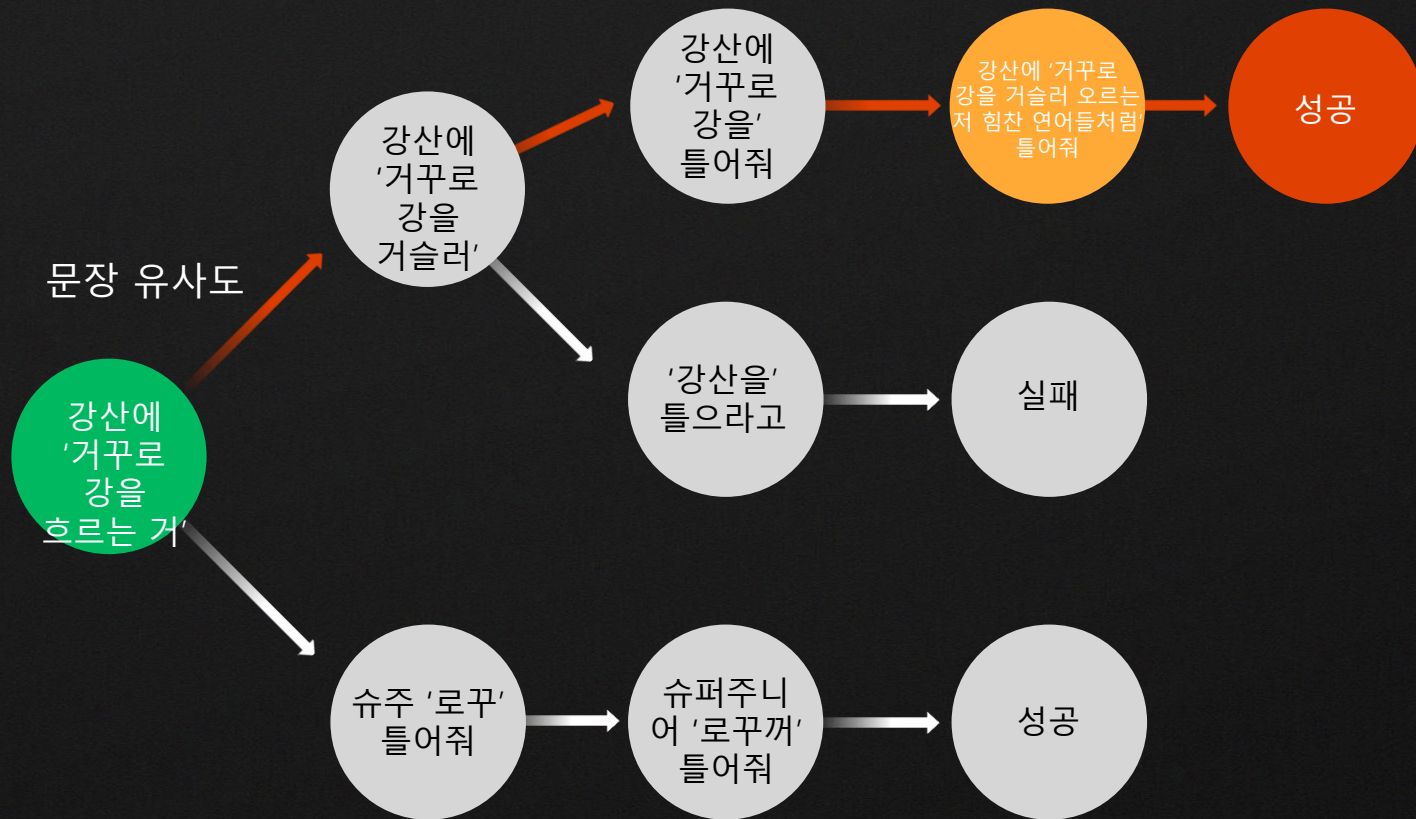


성공

# Query reformulation : Feedback-Based Self-Learning in Large-Scale Conversational AI Agents



# Query reformulation : Feedback-Based Self-Learning in Large-Scale Conversational AI Agents





## CHAPTER 3

# 인사이트

1

# GPT-3 와 대화 시스템

GPT-3

Generative Pre-trained Transformer



# Generative Pre-trained Transformer

방대한 데이터

언어 모델

제너럴리스트

# Generative Pre-trained Transformer

방대한 데이터

언어 모델

제너럴리스트

# Generative Pre-trained Transformer

방대한 데이터

언어 모델

제너럴리스트



# Generative Pre-trained Transformer

방대한 데이터


언어 모델

제너럴리스트

2

# Blender Bot 2.0

# Blender Bot 2.0



The diagram consists of three identical circular elements arranged horizontally. Each element is composed of three concentric circles. The innermost circle is a bright green color and contains white text. The middle and outer rings are a darker green color. The background of the entire slide is a dark charcoal gray.

Long Term  
Memory

Personalization

Internet Search



# Blender Bot 2.0




The diagram consists of three circular elements arranged horizontally. The first element on the left is a set of concentric circles with an orange center and brown outer rings. The middle element is a set of concentric circles with a green center and dark green outer rings. The third element on the right is a set of concentric circles with a green center and dark green outer rings. Each element contains text in its center.

Long Term  
Memory

Personalization

Internet Search

# Blender Bot 2.0



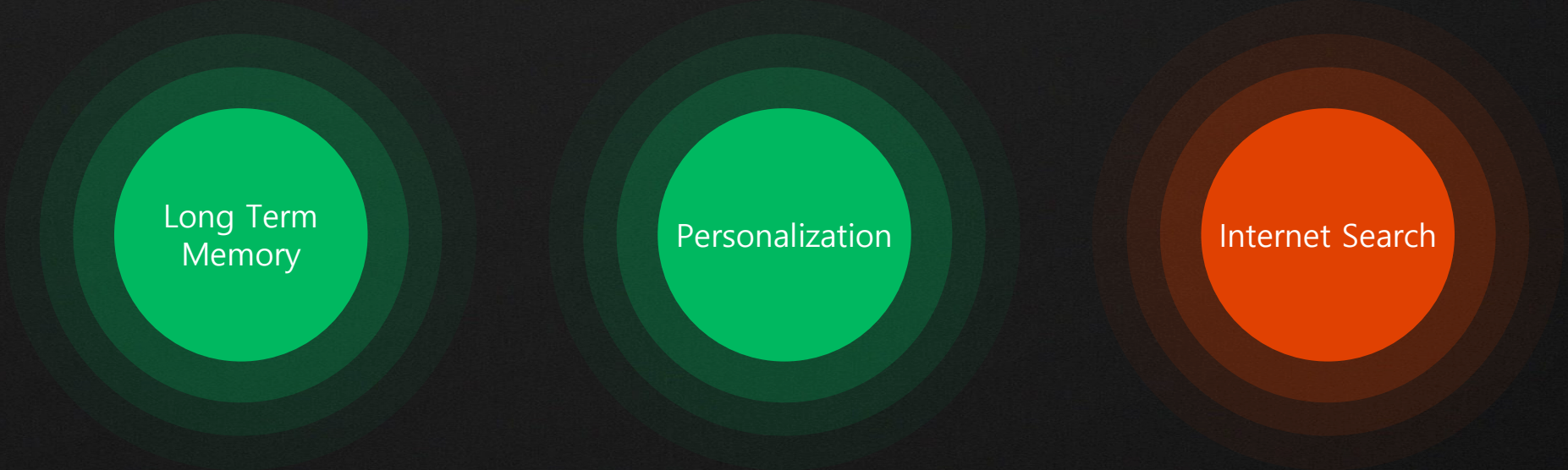
The diagram consists of three circular nodes arranged horizontally. Each node is composed of three concentric circles. The leftmost node has a bright green center and two darker green outer rings. The middle node has a bright orange center and two dark brown outer rings. The rightmost node has a bright green center and two darker green outer rings. The text is centered within each node's innermost circle.

Long Term  
Memory

Personalization

Internet Search

# Blender Bot 2.0



The diagram consists of three circular nodes arranged horizontally. The first two nodes, 'Long Term Memory' and 'Personalization', are green and have three concentric circles around them. The third node, 'Internet Search', is orange and has two concentric circles around it. All nodes are set against a dark gray background.

Long Term  
Memory

Personalization

Internet Search