Clova

Leveraging Search Results to make Conversational Agents Smarter

NLP Group
Inho Kang

NAVER Clova INSIDE

© NAVER Corp.

OVERVIEW - NAVER

Korea's largest Search Portal

74% of Search Market Share in South Korea
Search Engine + Various Online Services and Platforms

Knowledge-iN
Community QA platform. 400M Q&A database

NAVER Café, Blog and Post
2.3M blogs and 1M cafes
Daily 800K new user-generated documents

NAVER Dictionary and Encyclopedia
33+ languages, 3M knowledge entries

NAVER Shopping
700M+ products uploaded
from department stores to 'local' craft brands

OVERVIEW - LINE

LINE Global messenger widely spread to 230 countries

Monthly active users of 200 million

The most popular messaging application in Japan, Thailand, Taiwan, and more

Evolving from a mobile messenger to global mobile portal (LINE Pay, LINE News, …)

Clova : Cloud Virtual Assistant

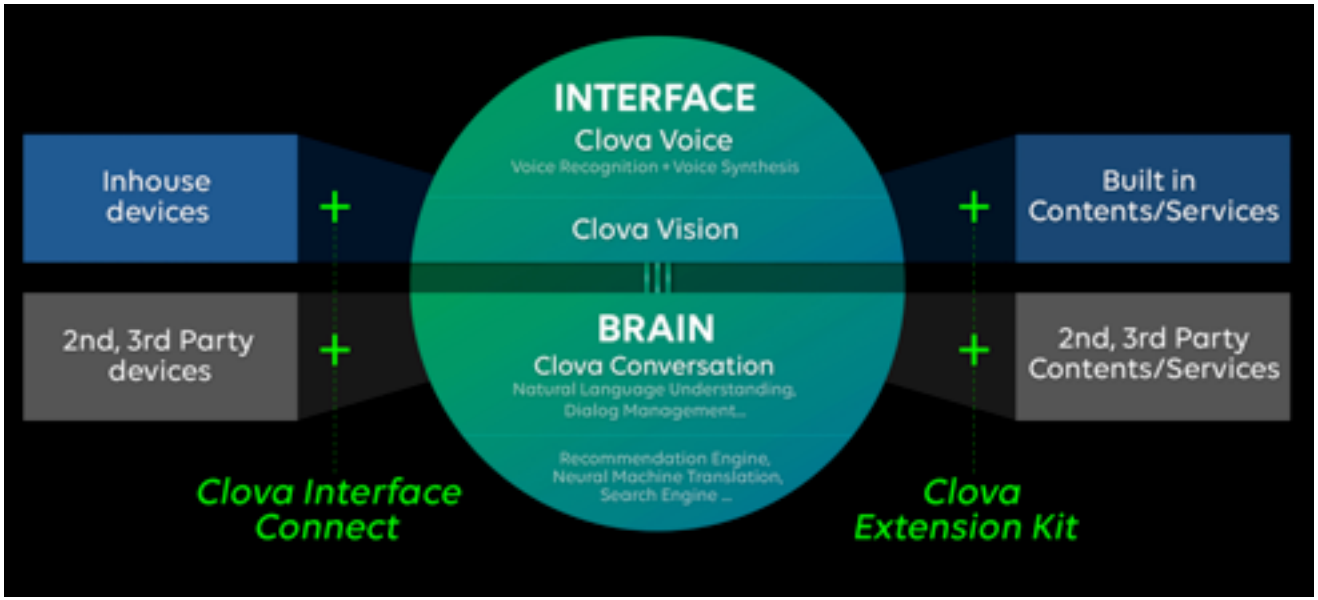
NAVER and LINE team up to develop a General-purpose AI platform

- Utilizes all contents accumulated by NAVER and LINE
- Makes use of all AI technologies that NAVER and LINE have been developed
- Synergies through Co-operation

General AI Platform

Goals:

Know you better
Get smarter everyday
Be with you everywhere



Clova Products

LINE's free voice-call function
Infrared home-appliance control

NAVER KOREA

Wave (2017)

Friends (2017)

Friends Minions (2018)

Away and AKI (2018)

On (2019)

LINE JAPAN

LINE JAPAN

Friends Doraemon (2018)

Clova Desk (2019)

Clova inside Products

Difficulties - Language Variation

Guide Sentence: **What did Tyrannosaurus eat?**

What do Tyrannosaurus eat?

What is Tyrannosaurus food?

What does the Tyrannosaurus eat?

How did the Tyrannosaurus get its food?

What does Tyrannosaurus eat to live

What kind of food do Tyrannosaurus eat?

What sort of food did Tyrannosaurus eat?

…

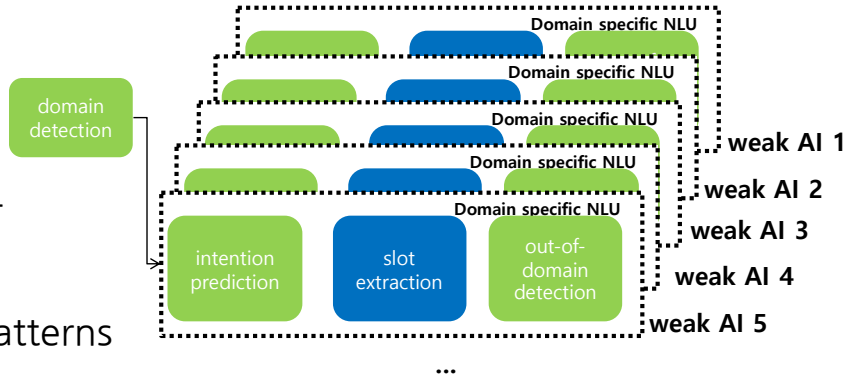
and many other interpretations ?

Implies simple patterns and user guidance are not enough

Language Understanding

Natural Language Understanding

- Finding Intentions and Parameters
- Analyzing 300M Search Queries / Day
- Language Variability Issues: Patterns + ML
- CNN-based Domain Detection
- 42 Major Domains / 347 Sub Domains
- 3,400 Properties, 4.8M Paraphrases for Patterns



inform:weather 〈address〉경기도 화성시〈/address〉 〈datetime="WE"〉주말〈/datetime〉 날씨

inform:weather 〈datetime season="PRESENT_REF" mod="APPROX"〉요즘〈/datetime〉 〈country〉라오스〈/country〉 기후는 어떤가요?

inform:musician 가수 〈name〉신용재〈/name〉가 있는 그룹 이름이 뭐였지?

convert:length 〈unit value="130" unit="INCH"〉130인치〈/unit〉가 〈unit name="METR"〉미터〈/unit〉로 하면 얼마지?

Choi, et. al, “Multi-labeled Domain Detection using CNN”, HCLT 2017

Query Rewriting : enhancing recall and accuracy

Utilizing query reformulations in the session logs

Lexically

Spelling correction

Semantically

Domain-aware

Relation-aware: Artist and Song, Album and Song

Context-aware

Location and Contacts: Near Places, Frequently contacted Friends

Implicit Verification

Did you mean?
Displaying related queries

Context-aware understanding

NAVER specialty

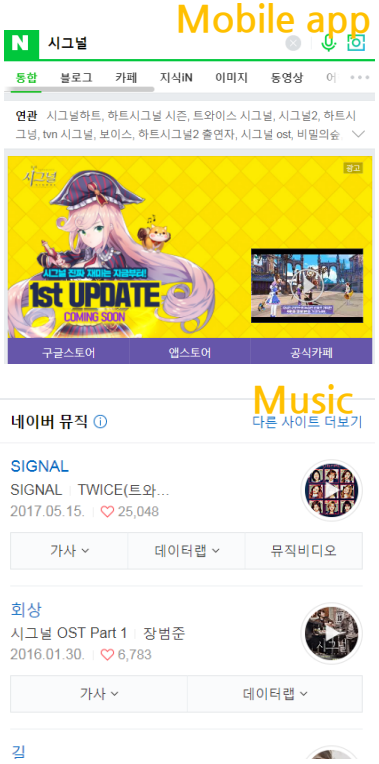
Mostly short and ambiguous queries

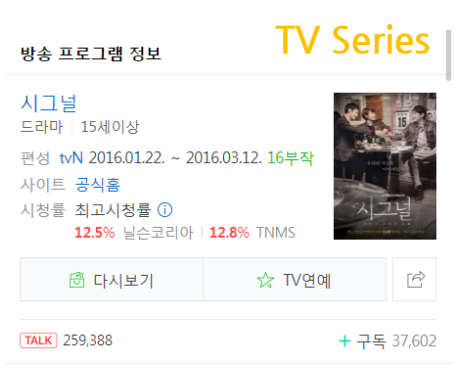
top 1, 2 queries of the smart set-top box are one word (Youtube, Porroro)

leveraging search users' preference to know most likely intention of a query

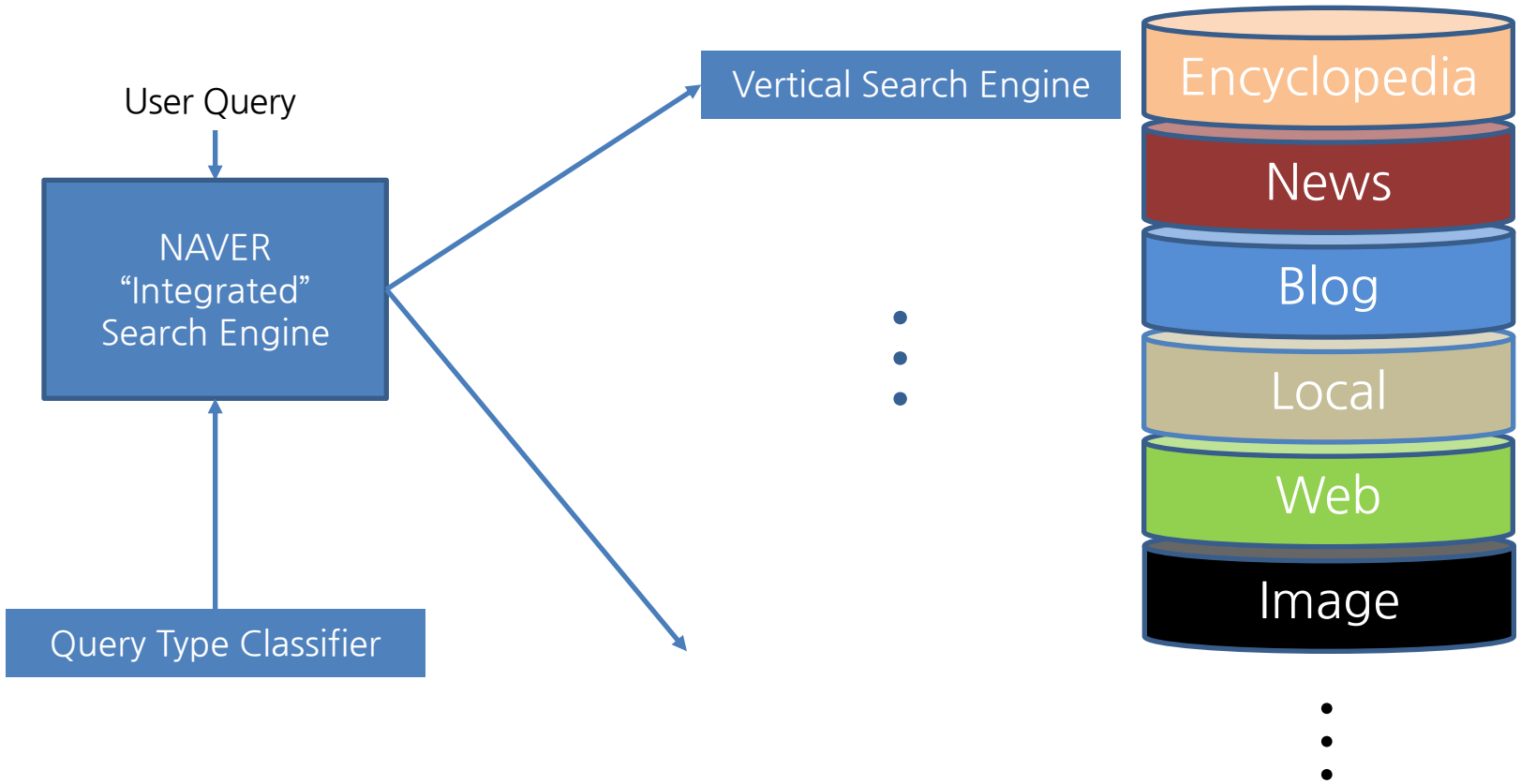
same domain/service but different object

Play 'Signal'





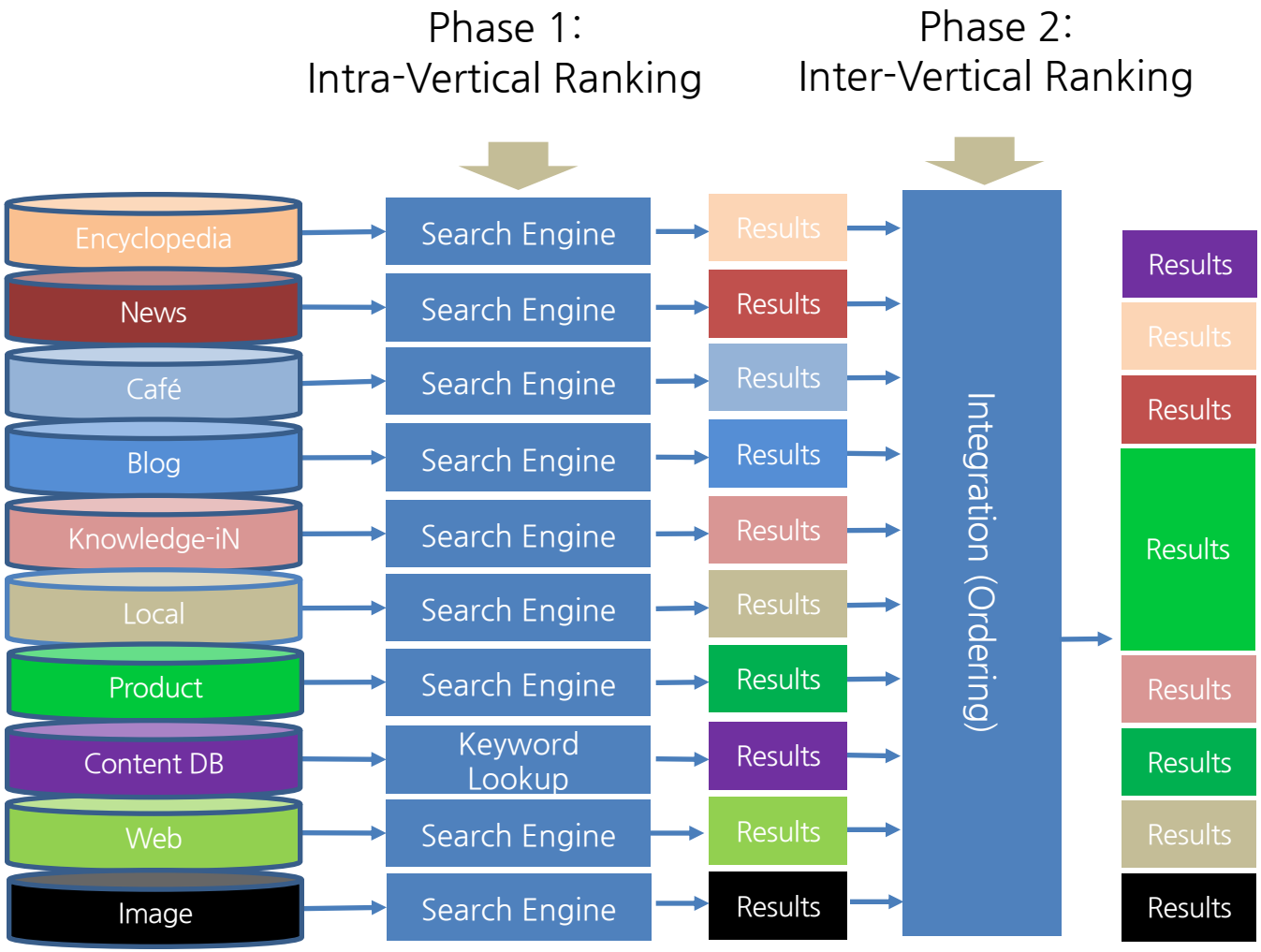
NAVER Search → Integrated Search



Recency (Steady, Trending, ⋯)
Navigational, Transactional, Informational
Query Category (Place, Shopping, ..)

NAVER Search: Two-Phase Ranking

Search results based user intention estimation

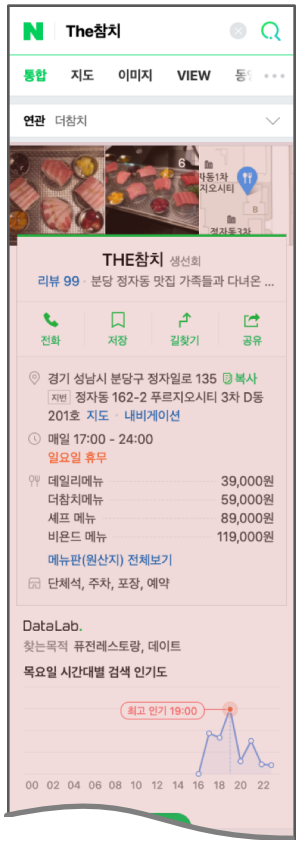


NAVER SERP Examples (Query: BTS)

Merging and Ranking, according to the top intention

NAVER SERP Examples (Query: TheTuna)

Merging and Ranking, according to the top intention

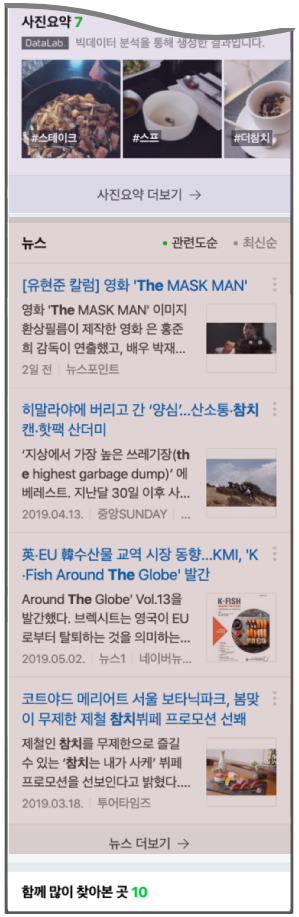


Local

Blogs

Images

News



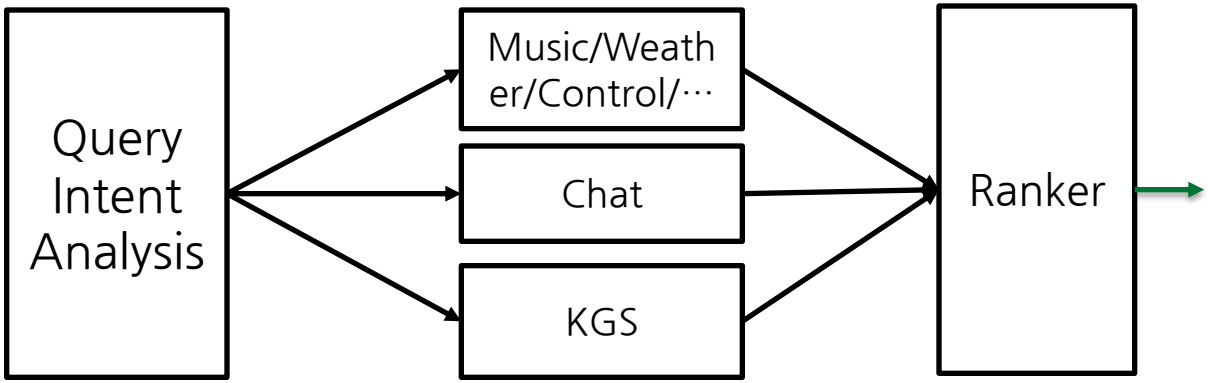
Web

Videos

Clova: Results based Prediction

Predicting a domain using sub-results of each domain

Especially, effective for Informational needs



マリオメーカー
마리오 메이커 (Mario Maker)

Hard to determine using only utterance itself

すみません、　わかりませんでした。
죄송합니다. 모르겠습니다. (I am sorry I have no idea)

ウィキペディアによると、『スーパーマリオメーカー』は、任天堂より2015年9月10日に発売されたWii U用 …
위키피디아에 따르면, “슈퍼마리오 메이커”는, 닌텐도로부터
(According to wikipedia, Super Mario is made by Nintendo…

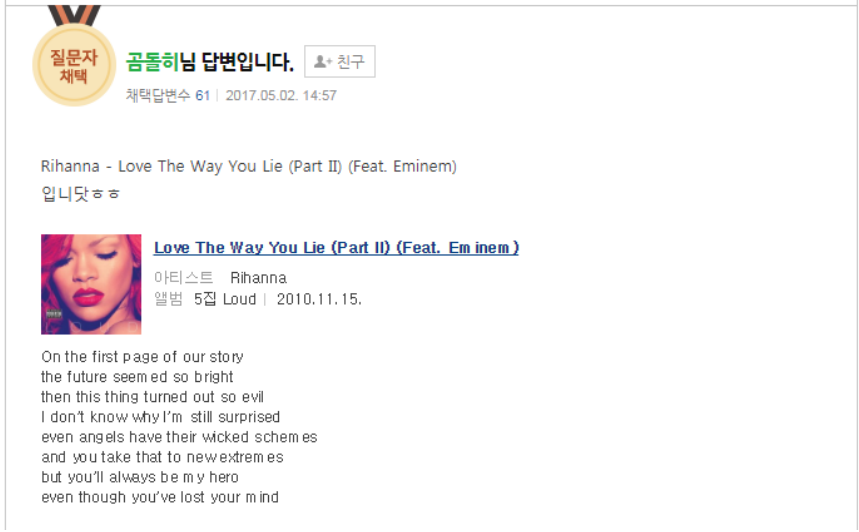
Knowledge-iN - platform to find good sources

“User-centric Q&A platform”

Anyone can share any kind of knowledge, know-how and tips.

What is the title of this song?

I came across this song.
The color of the album jacket is totally red and a woman is looking down.
The song starts like ‘u.u.hu.u.u.um … all the story…’
I remember the title ends with ‘ll’



Rihanna - Love The Way You Lie (Part II)

Clova Brain - Clova Q&A

KG-based Q&A

Wikipedia, Encyclopedia, News & Blog

cQA-based Q&A

FAQ matching with semantic matching

Knowledge-iN, and manually answered documents

Vertical Q&A

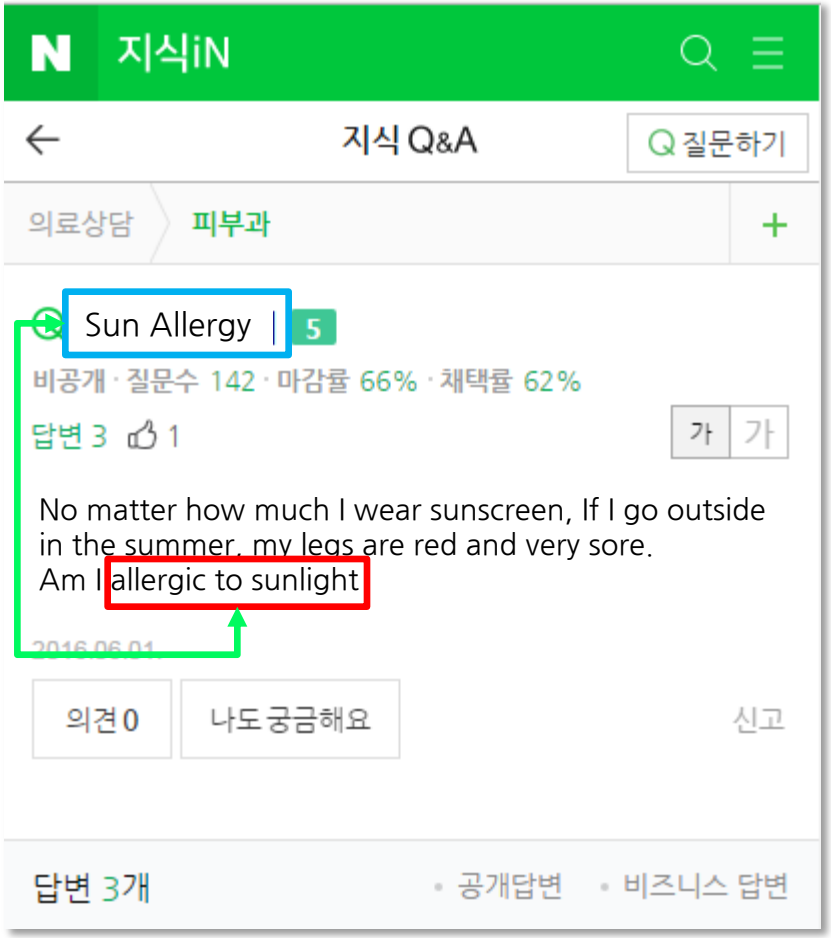
Movie/drama title and episode Q&A

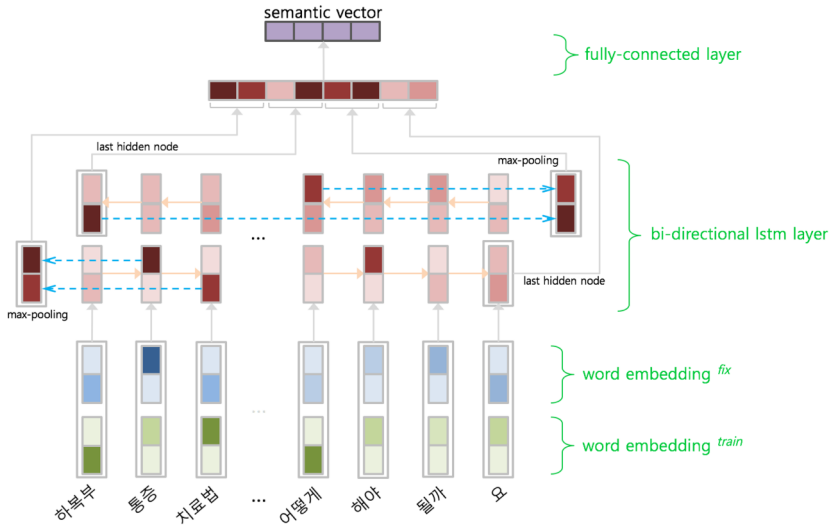
Language Understanding

NAVER specialty

Semantic based Matching

- Using 25M QA documents to embed a query to a vector
- Top score on Quora and SNLI dataset





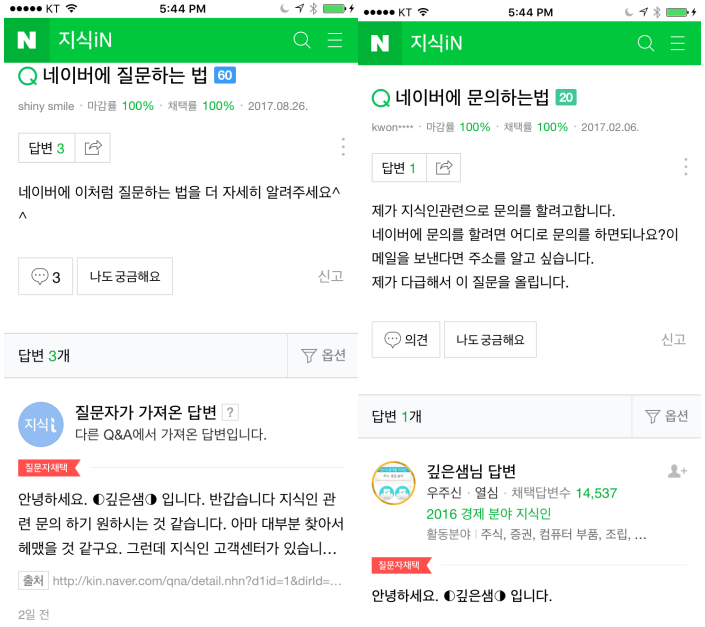
Kim, et. al, “Semantic Sentence Matching with Densely-connected Recurrent and Co-attentive Information”, AAAI 2019

Song, et. al, “Translation of Natural Language Query into Keyword Query Using a RNN Encode Decoder”, SIGIR 2017

Kim, et. al, “Question Retrieval using Deep Semantic Matching for Community Question Answering”, HCLT 2017

Question Embedding - cQA based Q&A

Providing previous QA
User-feedback based learning
Got the highest level in 1.5 year

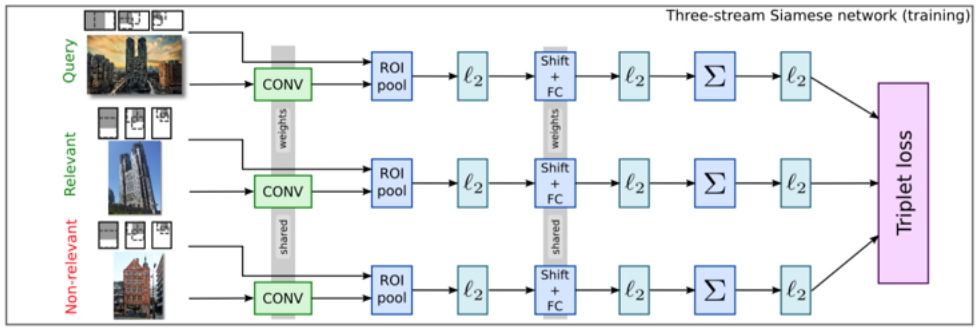


Search by Image

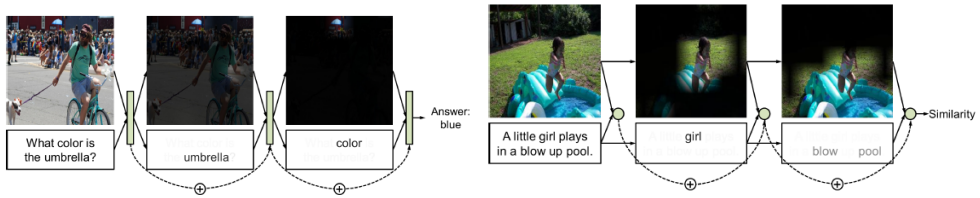
Want to know the breed of a cat

Computer Vision

Smart Lens: DNN-based Visual Search
Collaborating with NAVER Labs Europe (Fashion, Visual, Search, and so on)



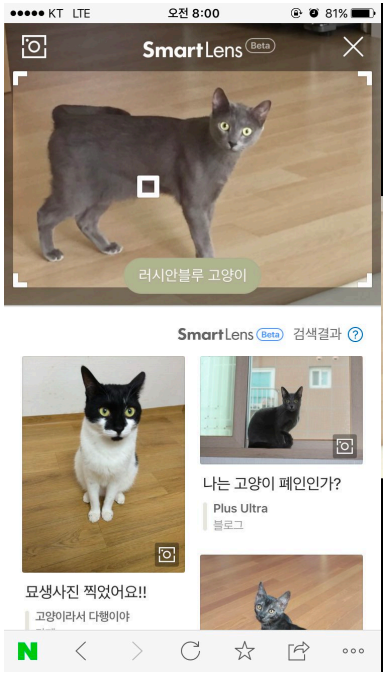
DNN-based Image Recognition Framework



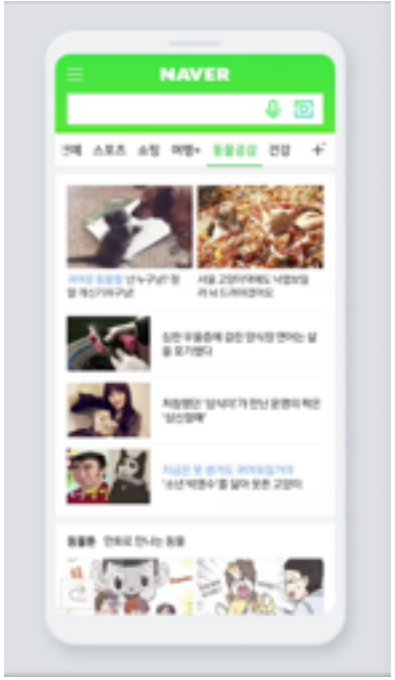
(a) DAN for multimodal reasoning. (r-DAN)

(b) DAN for multimodal matching. (m-DAN)

Clova & LABS Europe

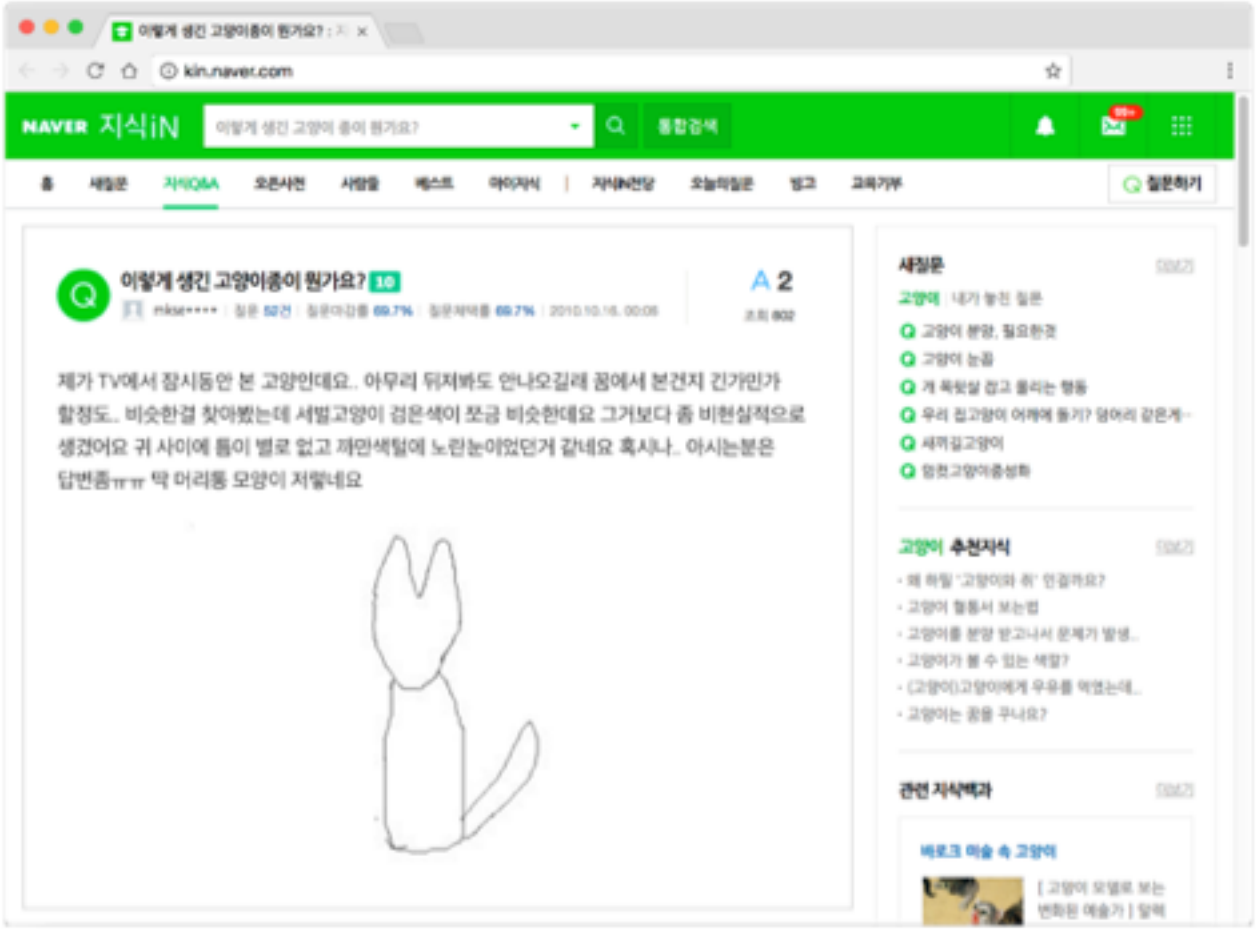


Smart lens



Shopping lens

Matching Images - Visual Question Answering



Not yet, but we are trying…

Domain-specific Q&A

NAVER specialty

Using relevant movie reviews
Current accuracy 85%

grandfather with dementia and robot 검색

main character is hitting a drum 검색

losing memory every 10 min. 검색

13 becomes 30 years old 검색

surviving on a lifeboat 검색

1. 라이프 오브 파이

2. 올 이즈 로스트

3. 캡틴 필립스

Communicating & Sharing

Share everything with your friends

Ask friends and experts directly and verbally

Leveraging your network

Wide range of services all over the world

People
location
Information
People

Influencer Search

LINE Integrated Search

LINE specialty

Communication Search

Chat

Open Chat

Official Account

Timeline

Contents Search

LINE NEWS

Weather

LINE Manga

LINE MUSIC

Celebrities

LINE Shopping

1 to 1 Platform
LINE ASK me

LINE ASK me
Search
Discover
Action

Talk Fortune

Over 1,000 fortunetellers

Talk Care

Over 200 experts in diverse fields

(mental care,career,business,love)

+1

Concluding Remarks

Come and Grow with us

We are expanding fast with offices in Korea, Japan, France, HongKong and more happening now

https://clova.ai/en/research/careers.html

-

End of Document

-

Thank You.

-

NAVER

© NAVER Corp.