



THUIR-LL at the NTCIR-16 Lifelog-4 Task: Enhanced Interactive Lifelog Search Engine

**Zhiyu He, Jiayu Li, Wenjing Wu,
Min Zhang*, Yiqun Liu, Shaoping Ma
Tsinghua University
z-m@tsinghua.edu.cn**

Outline



1. Lifelog-4 Task

2. Interactive Search Engine

- Category-Based Query-text Parsing
- New Feedback Mechanism
- User Attention Inspired Result Presentation for Interaction

3. Experiment and Result

4. Conclusion and Future Work



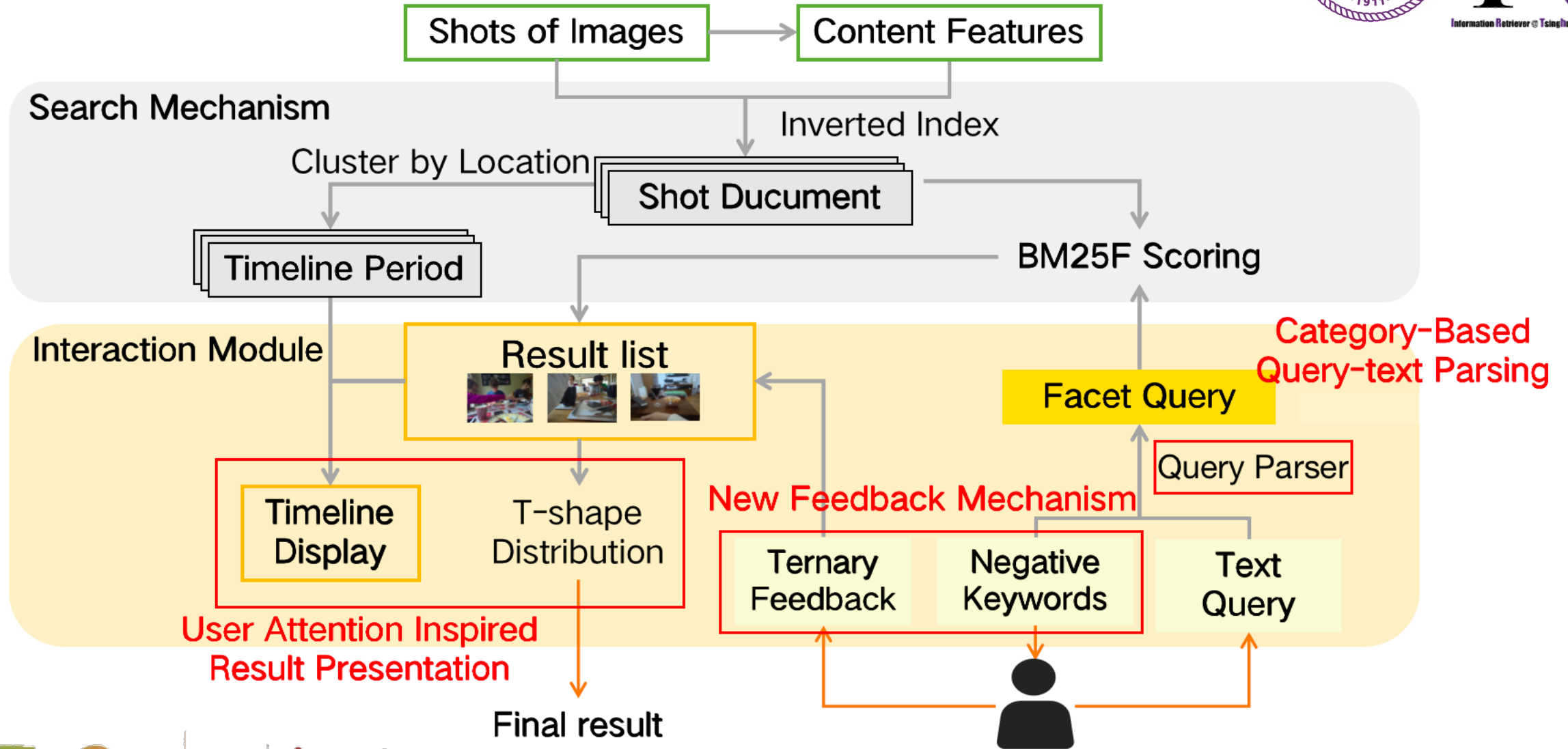
Lifelog-4 Task



- **Retrieve a number of specific moments in a lifelogger's life.**
 - Input: query of a topic
 - Output: relevant lifelog images
 - Interactive Manner: the user involved in the search process to provide relevance feedback or query reformulation.
- **Dataset:** 114 days, 183,432 wearable camera images with extracted visual concepts and textual metadata.



Framework of Interactive Search Engine



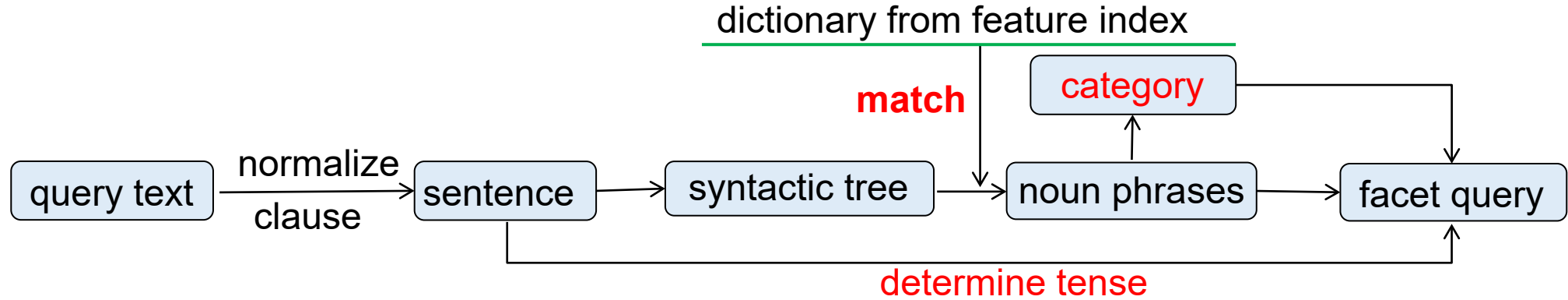


Interactive Methods

- Category-based Query-text Parsing
 - Considering verb's tense and noun's category
- New Feedback Mechanism
 - Ternary feedback for results
 - Field query with negative key words
- User Attention Inspired Result Presentation for Interaction
 - T-shape distribution
 - Timeline viewing function



Category-based Query-text Parsing

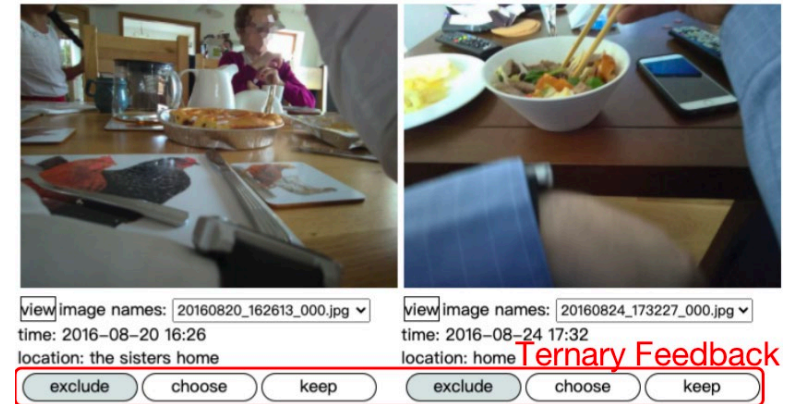
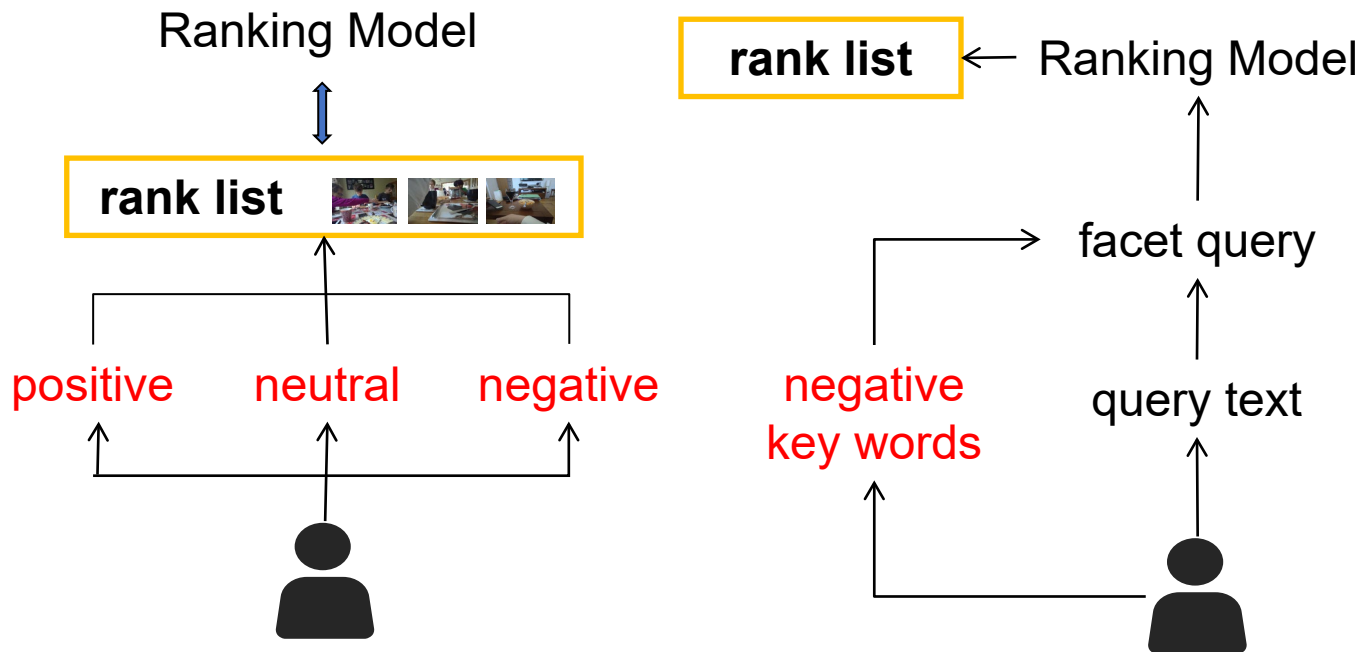


Query Text	Facet Query
I was lifelogsed when eating lunch at work in my office.	tags in this shot: [food, office]; approximate time: [noon]; location in this shot: [office, work]
Find examples of when I was in meetings with other people in rooms with red carpet. Before that, I had lived in hotel.	tags in this shot: [red carpet, room, rooms, people]; tags in past 1 hour : [hotel]; location in past 1 hour : [hotel]



New Feedback Mechanism

- Ternary feedback for results
- Field query with negative key words



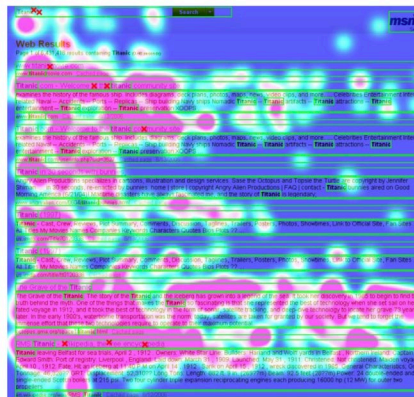
The screenshot shows a 'Negative Feedback' section with three input fields. The first field is labeled 'Exclude Images with tag:', the second 'Exclude Images with location:', and the third 'Exclude Images with activity:'. Each field has a small 'x' icon in the top right corner.

Negative Feedback Boxes



User Attention Inspired Result Presentation for Interaction

- Users' examination behavior for text result page and image result page may be different.
- Text result page: "F-shaped distribution" [1]
- Image result page: "T-shaped distribution"[2]



Heat map of the fixation
for text result page

0.831	1.036	0.888	0.890	0.671
0.505	0.697	0.694	0.653	0.560
0.381	0.381	0.591	0.479	0.326
0.337	0.442	0.509	0.478	0.302
0.314	0.356	0.442	0.392	0.322
0.316	0.343	0.380	0.343	0.272
0.300	0.346	0.375	0.327	0.225
0.230	0.296	0.389	0.373	0.237
0.258	0.348	0.321	0.426	0.255
0.251	0.261	0.341	0.289	0.184

Distribution of examination
duration in image search

[1] Edward Cutrell and Zhiwei Guan. 2007. What are you looking for? an eye-tracking study of information usage in web search. CHI '07.

[2] Xiaohui Xie, Jiaxin Mao, Yiqun Liu, et al. 2019. Grid-based Evaluation Metrics for Web Image Search. IWWW '19



User Attention Inspired Result Presentation for Interaction



Timeline Viewing Button

finding an example when i was chopping vegetables into a bowl

Fill in Forms

Detail Description

Time

Date:

Weekday:

Approximate time:

Range of hour:

Objects

Objects in the scene:

Objects seen in future 1 hour:

Objects seen in past 1 hour:

Biometrics Logs

Total Elevation is

Total calorie is

Total Step is

Average speed is

Average heart rate is

Semantic Description

Description in this shot

Location:

Activity:

Description for future 1 hour

Location:

Activity:

Description for past 1 hour

Location:

Activity:

topic id:

Negative Feedback

Exclude Images with tag:

Exclude Images with location:

Exclude Images with activity:

Negative Feedback Boxes

Result Page

LIFE RECORD

Home Search by Text Search by Image Timeline

Find examples of when I use of home being medication at the weekend. The medication or supplements were clearly visible in my hand.

Detail Description

Date:

Weekday:

Approximate time:

Range of hour:

Objects

Objects in the scene:

Objects seen in future 1 hour:

Objects seen in past 1 hour:

Biometrics Logs

Total Elevation is

Total calorie is

Total Step is

Average speed is

Average heart rate is

Semantic Description

Description in this shot

Location:

Activity:

Description for future 1 hour

Location:

Activity:

Description for past 1 hour

Location:

Activity:

topic id:

Negative Feedback

Exclude Images with tag:

Exclude Images with location:

Exclude Images with activity:

Timeline Viewing





Experiment and Result

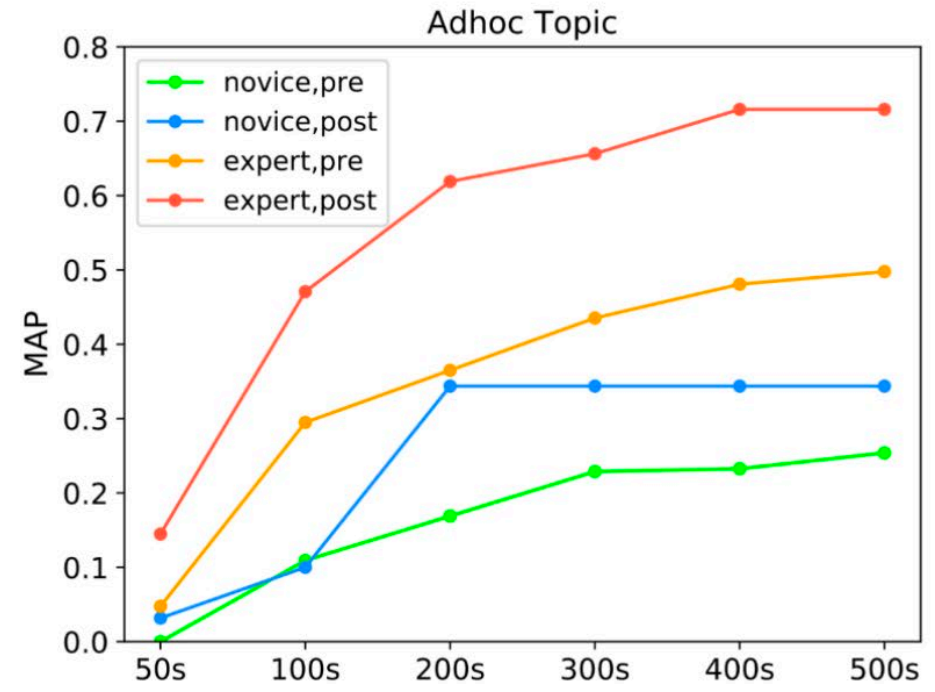
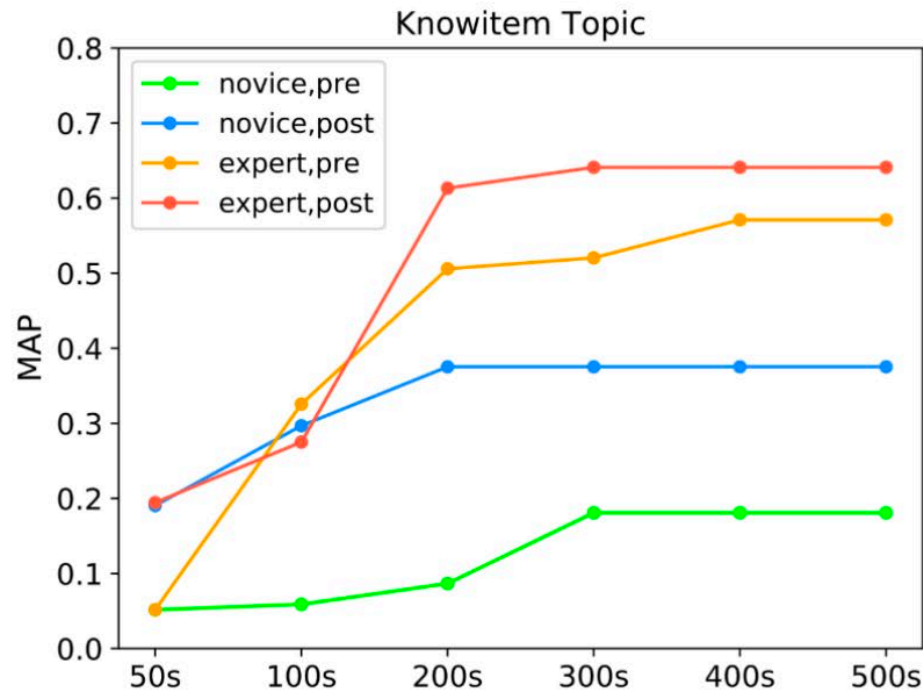
- Post vs. pre: the enhancement has improved the system in various metrics
- Expert vs. novice: user's familiarity has a significant influence on the effects
- Adhoc vs. knownitm: retrieval of multi-scene topics is more likely to be affected by differences in user understanding

	Knownitem topic				Adhoc topic			
	novice		expert		novice		expert	
	pre	post	pre	post	pre	post	pre	post
num_ret	24	26	37	40	215	233	159	162
num_rel_ret	19	21	31	36	51	66	118	103
map	0.2086	0.3753	0.5711	0.6410	0.3047	0.4857	0.4850	0.7160
Rprec	0.2086	0.3753	0.5842	0.6410	0.2556	0.5220	0.5487	0.7571
recip_rank	0.6667	0.8333	1.0000	1.0000	0.6848	0.6917	0.5931	0.8431
P_5	0.3333	0.4000	0.5667	0.6000	0.5600	0.6400	0.5000	0.8000
P_20	0.1583	0.1750	0.2583	0.3000	0.4300	0.4300	0.4917	0.6000



Experiment and Result

- The interactive method improves significantly in the later stage while it may not be effective in the early stage.



MAP at various time cut-offs





Conclusion and Future Work

- Interactive Search Engine
 - Category-Based Query-text Parsing
 - New Feedback Mechanism
 - User Attention Inspired Result Presentation for Interaction
- Future Work
 - Optimizing methods to eliminate noise in the dataset
 - Exploring how the user's attention distribution is in lifelog interactive search scenarios





Information Retriever © Tsinghua University

THANKS FOR YOUR ATTENTION !

Zhiyu He, Jiayu Li, Wenjing Wu,
Min Zhang*, Yiqun Liu, Shaoping Ma
Tsinghua University
z-m@tsinghua.edu.cn

