The NTCIR-11 **IMine** Task Kickoff Meeting

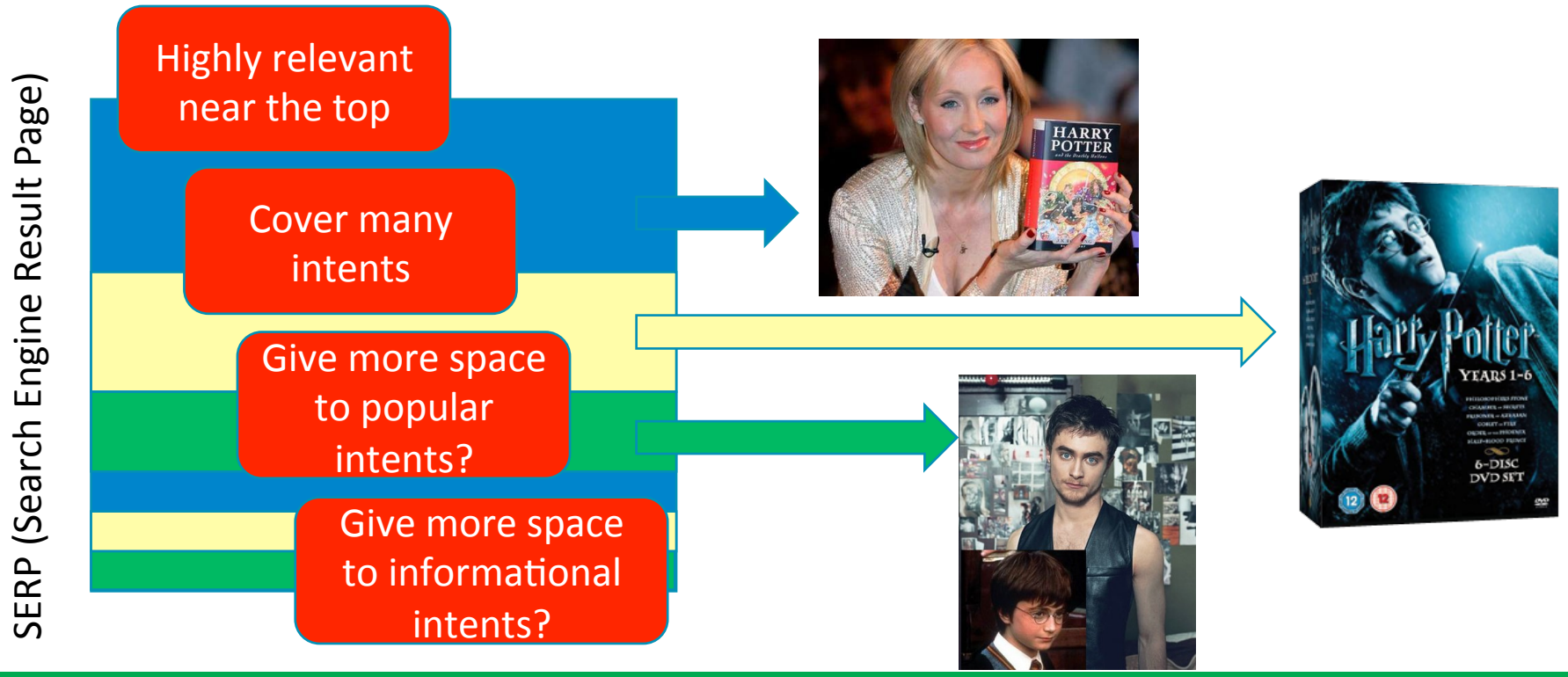
Yiqun Liu, Ruihua Song, Min Zhang, Zhicheng Dou,
Takehiro Yamamoto, Makoto P. Kato, Hiroaki Ohshima,
Ke Zhou

http://www.thuir.org/IMine/

September, 2013@NII, Tokyo

Diversified search

- Given an ambiguous/underspecified query, produce a single Search Engine Result Page that satisfies different user intents!
- Challenge: balancing relevance and diversity



IMine

Understanding user **intents** in Web search

Intent Mining

曖昧

（ambiguous）

The IMine task

Three subtasks

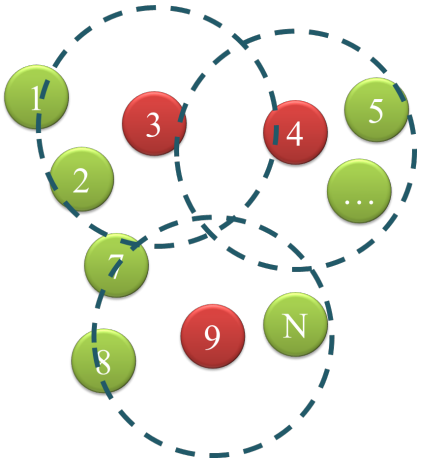
- **Subtopic Mining (SM)**: **Chinese, English, Japanese**
  INPUT : query (e.g. “harry potter”)
  OUTPUT: ranked list of subtopic string
  (e.g. “harry potter book, harry potter film, harry potter the character…”)
- **Document Ranking (DR)**: **Chinese, English**
  INPUT: query (e.g. “harry potter”)
  OUTPUT: diversified ranked list of web pages
- **Search Task Mining (TASKMINE)**: **Japanese**
  INPUT : query (e.g. “pollen allergy treatment ”)
  OUTPUT: ranked list of task string which satisfies the given query
  （e.g: “Laser surgery”， ”antiallergic drug”， ”allergy mask”）

More User Behavior Data

- THUIR@INTENT/INTENT2: More user behavior data led to better performance.
- More raw click-through data: SogouQ has doubled its size to include click-through data collected from Sogou.com in 2011
  - 1.85GB => 3.85GB, over 40M user clicks
- More subtopic candidates generated from more recent user behavior data.
  - Search engine data provided by two major Chinese search engines will be adopted
- Find out whether more logs help improve SM/DR performance

Intent Annotation Using Logs

- SM results in pools will be clustered with click-through/pseudo RF data at first to generate preliminary candidate intent groups.
- Query frequency information will be taken into consideration during subtopic importance voting process.
- Data source: recently collected data from Sogou for Chinese SM, Bing for English/Japanese SM



- More credible SM qrels with less annotation efforts
- Perhaps the reusability of results could be increased

Crowd-Sourcing based Evaluation

- A search-engine-like annotation interface to collect feedback information from a relatively large number of unprofessional users (e.g. 50+ undergraduate students)
- Data collected from the interface: query, click, examinatio $P(C_i = 1) = P(E_i = 1)P(R_i = 1)$
- High correlation with examination
  - Preliminary results on 10 users.
  - KAPPA: 0.65, Accuracy: 0.83.
- Find out whether D# measures accords with user satisfaction

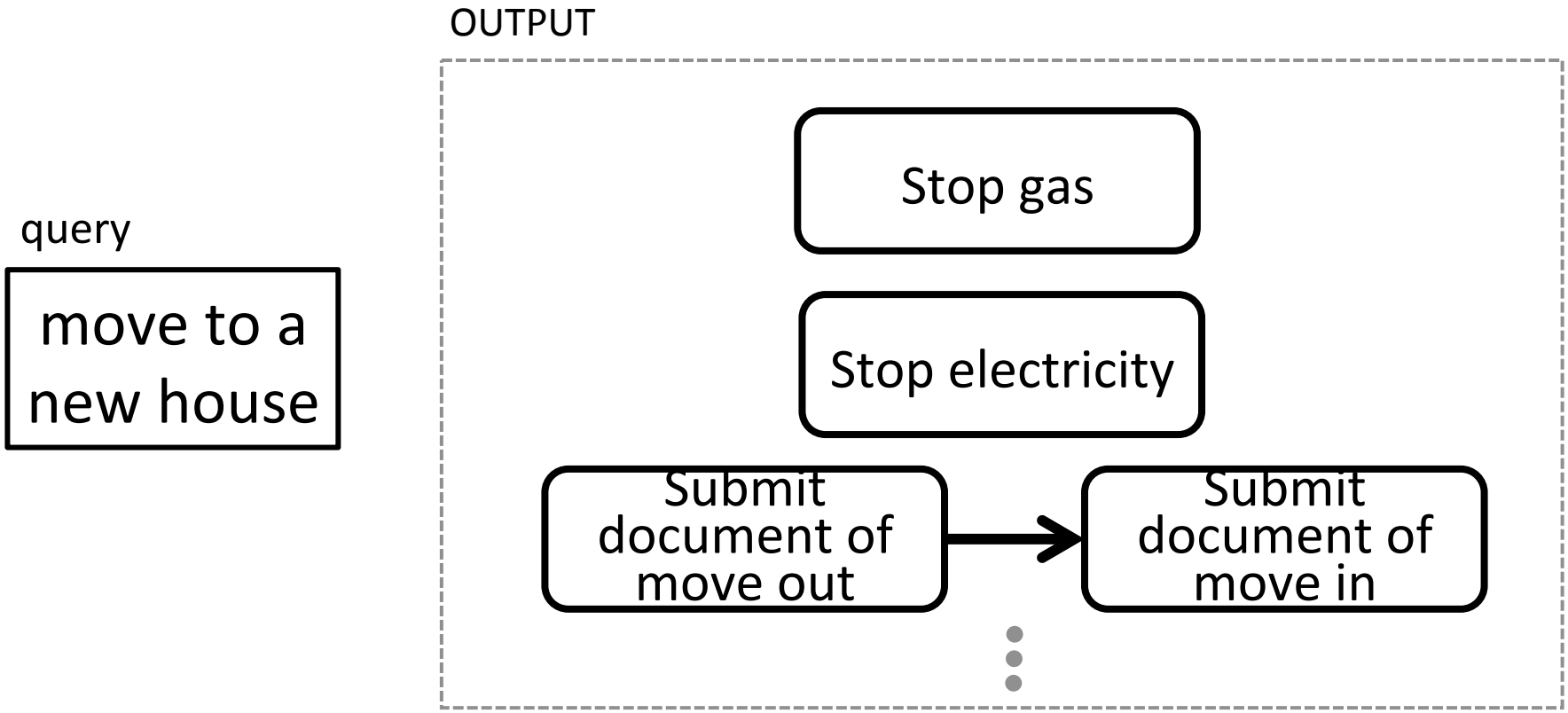
INTENT with/for Knowledge Graph

- Fuji in Wikipedia:
  - Fuji the actress
  - Fuji the Mountain
  - Fuji the apple

- Ambiguous queries: NTCIR INTENT #0205: 功夫(kung fu)，#0206: 生日快樂(happy birthday)
- Mining and evaluating hierarchical user intents
  - SM task: return a two-level hierarchical list of subtopics
  - First level: at most 10 major categories of user intent (“Microsoft windows” for “windows”)
  - Second level: other minor subtopics under the categories (“windows update”, “windows 8.1 installation”)

Task Mining (TASKMINE) subtask

Find tasks that satisfies the given query



Organisers: Takehiro Yamamoto, Makoto Kato, Hiroaki Ohshima

http://www.dl.kuis.kyoto-u.ac.jp/ntcir-11/taskmine/

Summary of IMine

	INTENT2	***IMINE***
Number of Topics	• Chinese: 100 • Japanese: 100 • English: 50	• Chinese: 50 • Japanese: 50 • English: 50
DR task setting	• Chinese: SogouT (Ver.2008) • Japanese: ClueWeb JA	• Chinese: SogouT (Ver.2008) • English: ClueWeb12-B13
Manual annotation efforts	• SM: 100 Chinese topics, 50 English topics, 100 Japanese topics • DR: 100 Chinese topics, 100 Japanese topics, pool depth=20	• SM: 50 Chinese topics, 50 English topics, 50 Japanese topics • DR: 50 Chinese topics, 50 English topics, 50 Japanese topics, pool depth=20
Support from log analysis for annotation	No	Support from log analysis for SM/DR annotation
Crowd sourcing	No	Crowd sourcing for Chinese DR
Subtopic candidate	Query suggestions from Bing, Google, Sogou and Baidu	• Query suggestions from Bing, Google, Sogou, Yahoo! and Baidu • Query facets generated by MSR from search engine results • Query facets generated by THU from Sogou log data
User behavior data	SogouQ (data collected in 2008): 2GB approximately	SogouQ (data collected in 2008 and 2011): 4GB approximately
DR Baseline	• Chinese DR baseline • Japanese DR baseline	• ClueWeb12-B13 retrieval service is provided by CMU • SogouT retrieval service is provided by Tsinghua

Thank you

http://www.thuir.org/IMine/

http://www.dl.kuis.kyoto-u.ac.jp/ntcir-11/taskmine/