NTCIR-18 Kickoff Event

Transfer-2 Task

Hideo Joho  Atsushi Keyaki  Yuuki Tachioka  Shuhei Yamamoto
University of Tsukuba  Hitotsubashi University  Denso IT Laboratory  University of Tsukuba
Table of contents

● Overview of Transfer Task
  ○ Summary of Transfer-1

● Transfer-2 Subtasks
  ○ Dense Cross-Language Retrieval (DCLR)
  ○ Dense Multimodal Retrieval (DMR)
  ○ Retrieval Augmented Generation (RAG)

● Tentative Schedule

● Contact
Overview of Transfer Task  

(in a broad sense of model application)

Existing Resources
LLM/Ranker/Reranker
Query Generator
Relevance Assessor
Web/English

Language
Domain
Task
Modality
Ranking

New IR Applications
with Dense Vectors
in Japanese
Summary of Transfer-1

- **Subtasks**
  - Dense First Stage Retrieval
    - Adhoc retrieval with NTCIR-1/2 test collections
    - Japanese academic corpus
    - nDCG@1000
  - Dense Reranking
    - Top 1000 from BM25 to submit top 100
    - nDCG@20, MRR

- **Participation**
  - 3 groups + organizer group
  - 29 runs for first stage, 25 runs for reranking

- See the [overview paper](#) for details
Subtask 1: Dense Cross-Language Retrieval (DCLR)

- Japanese to/from other languages
- Adhoc search / Reranking
- Currently in discussion with TREC NeuCLIR organizers
  - 2023 collections: Chinese, Persian, Russian (2016 to 2021)
  - Domain: News
  - Queries: 54~60 (the above three languages + English)
  - Metrics: nDCG@20 and others
Subtask 2: Dense Multimodal Retrieval (DMR)
Subtask 3: Retrieval Augmented Generation (RAG)

Question and answering task (SOTA is a two-stage system)

- **First stage (document retrieval from given dataset)**
  - LLM-based retrieval or conventional methods
  - Retrieval performance is measured

- **Second stage (answer generation by using LLM with RAG)**
  - Two types of RAGs for LLM
    - Corpus-based RAG (prompting by using relevant documents)
    - Data-based RAG (prompting by using any data such as triples of KG)
  - Answer generation performance is measured

- **RQ**: Does the performance of answer generation depend on the retrieval performance?
<table>
<thead>
<tr>
<th>Date</th>
<th>Activity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mar 29, 2024</td>
<td>Kick-Off Event</td>
</tr>
<tr>
<td>Jun 1, 2024</td>
<td>Task guidelines and dataset release (might vary across subtasks)</td>
</tr>
<tr>
<td>Nov 1, 2024</td>
<td>Task registration due</td>
</tr>
<tr>
<td>Dec 1, 2025</td>
<td>Run submission due (might vary across subtasks)</td>
</tr>
<tr>
<td>Feb 1, 2025</td>
<td>Evaluation results return // Overview paper (draft) release</td>
</tr>
<tr>
<td>Mar 1, 2025</td>
<td>Participant paper (draft) due</td>
</tr>
<tr>
<td>May 1, 2025</td>
<td>Camera-ready due for participant papers</td>
</tr>
<tr>
<td>Jun 10-13, 2025</td>
<td>NTCIR-18 Conference</td>
</tr>
</tbody>
</table>
Contact

- URL: https://github.com/ntcirtransfer/transfer2/discussions
- Email: ntcir-transfer-contact@googlegroups.com
- X (Twitter): #ntcir_transfer
- Slack: Registered participants only