

From Shared Tasks to Benchmark Datasets: Lessons from NTCIR's 26-Year Journey

Chung-Chi Chen¹, Makoto P. Kato^{1,2}, Noriko Kando¹

¹National Institute of Informatics, Japan

²University of Tsukuba, Japan



<http://lrec.nlpfin.com/>

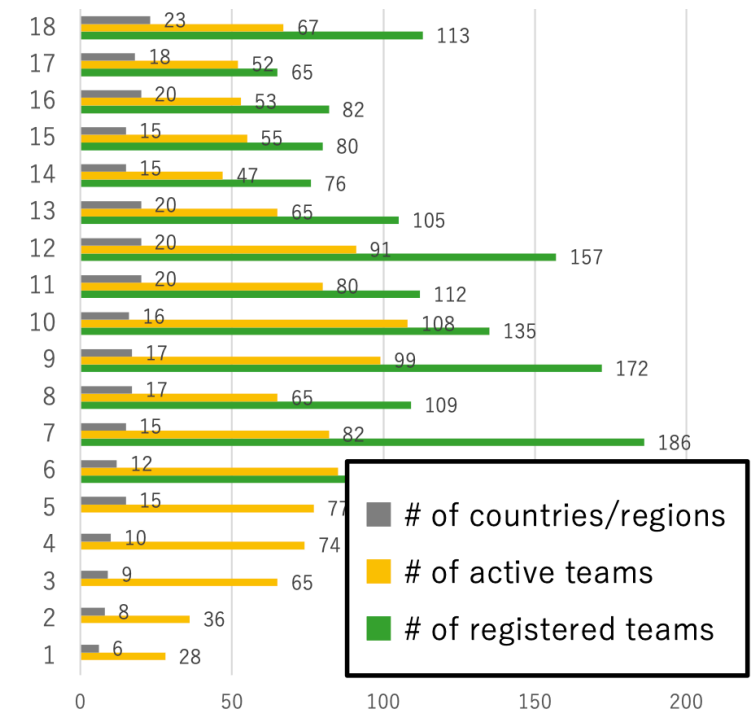


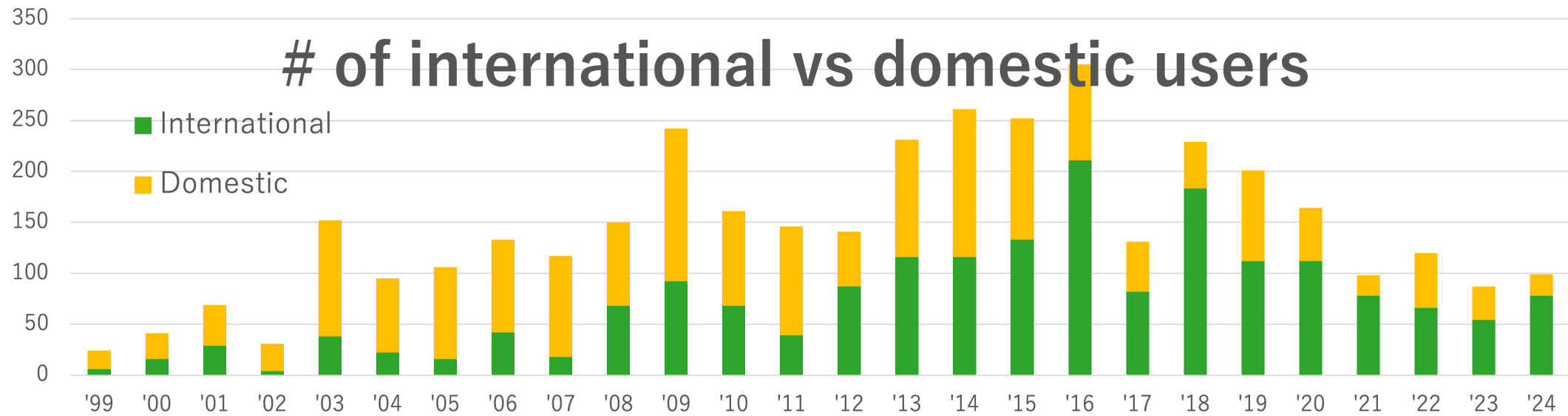
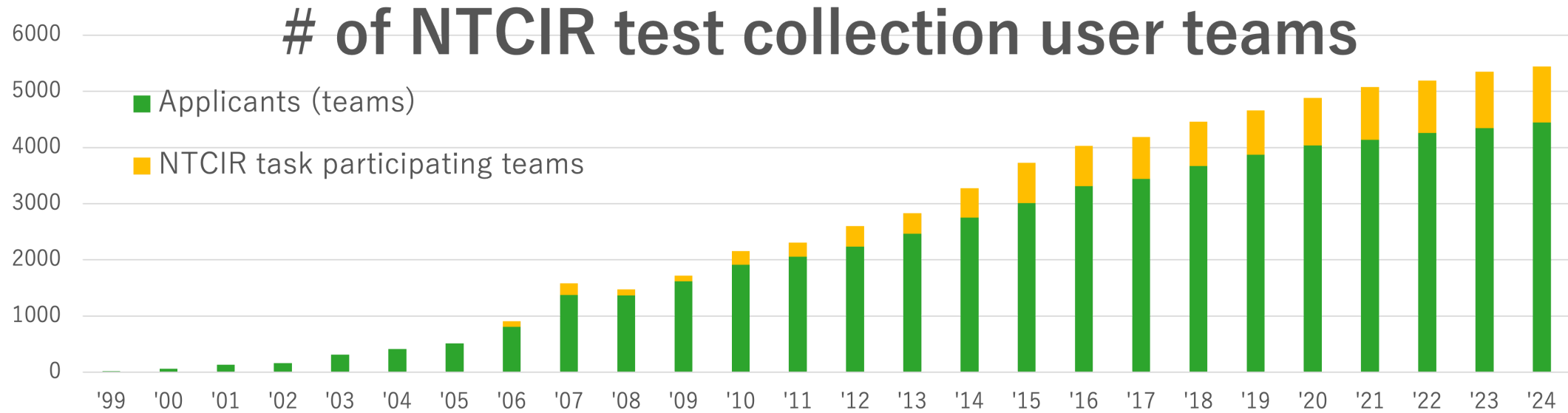
NTCIR (NII Testbeds and Community for Information access Research)

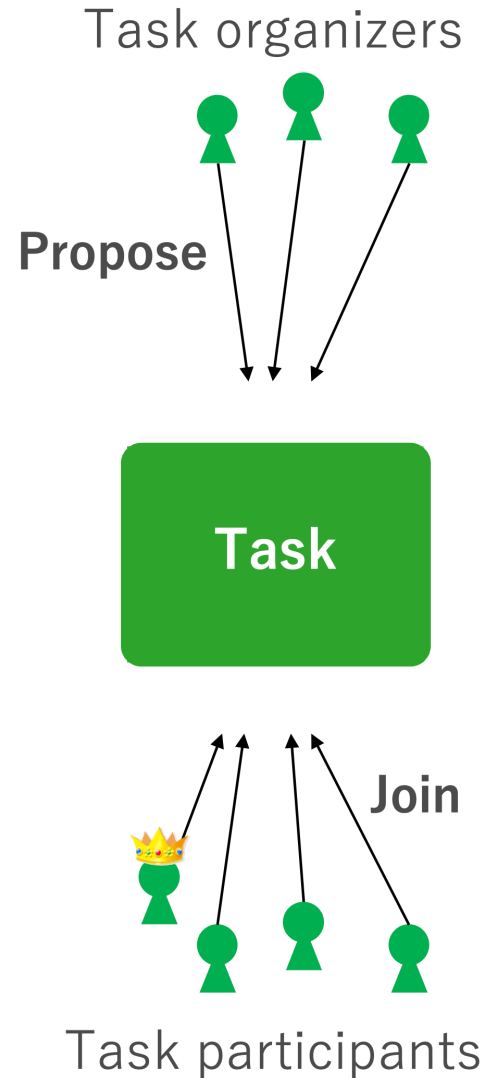
A series of evaluation campaigns designed to enhance research in information access technologies by providing an evaluation infrastructure: datasets, evaluation methodologies, and forum

- **Project started in late 1997**
 - 18 months cycle
- **Test collections**
 - Scientific, news, patents, web, CQA, Wikipedia, exams, spoken, microblogs, lifelogs, financial and medical documents
 - Chinese, Korean, Japanese, and English
- **Tasks (Research areas)**
 - IR: cross-lingual, patents, web, geo, time, math, spoken, recipe, microblogs, lifelogs, open data, session search
 - QA: monolingual, cross-lingual, interactive, IR for QA, complex, exams
 - Summarization, trend info., patent maps, inference, 2-layer summarization for mobile, dialogue generation and evaluation
 - Opinion analysis, text mining, intent mining, link discovery, information extraction, action detection, conversation, reading comprehension, etc.

of participants in NTCIR







1. Set up a task of their interest
2. Build a dataset
3. Have participants develop systems
4. Evaluate the participants' systems
5. Gain insights and write papers

WIN-WIN

Organizers: collaboratively tackle a problem

Participants: focus on solving the problem

1. Choose a task of interest
2. Download the dataset and develop a system that meets the task requirements
3. Have the system evaluated by the task organizers
4. Gain insights and write papers

- NII is a **Japanese** national research institute for informatics, located in Tokyo
 - Formerly National Center for Scientific Information System (NACSIS)
- NII has been hosting NTCIR since 1997, providing the infrastructure, resources, and organizational support necessary for the campaign's success
- While NII is the primary host, NTCIR has been supported by many researchers and institutions **worldwide**, who have participated in organizing tasks, providing data, and submitting runs



- **Noriko Kando**

- Professor emeritus and specially appointed professor at NII, honorary chair of NTCIR, and **the founder of NTCIR**
- Research interests: information retrieval, IR evaluation methodologies, interactive IR, and search-as-learning



- **Makoto P. Kato**

- Associate professor at the University of Tsukuba and NII, and general co-chair of NTCIR
- Research interests: information retrieval, IR evaluation methodologies, and retrieval algorithms



- **Chung-Chi Chen**

- Assistant professor at NII, and program committee co-chair of NTCIR
- Research interests: Agentic Framework Design, Human-Agent Teaming, High-Stakes Decision-Making, and Responsible Human-Agent Society



- **Part 1: NTCIR History and Test Collection Development (Noriko Kando)**
 - The origins of NTCIR, its evolution, key milestones in NTCIR's history, and community development.
- **Part 2: NTCIR Participants, Experiments, and Key Findings (Makoto P. Kato)**
 - A deep dive into the breadth of NTCIR tasks, spotlight on co-organized tasks, lessons learned from a variety of tasks, and a comparison with other attempts.
- **Part 3: Organizing Evaluation Campaigns and Dataset Sharing (Chung-Chi Chen)**
 - An exploration of the practical aspects of organizing evaluation campaigns, including task proposal development, dataset collection and curation, coordination with teams, and dissemination of results and resources.