Overview of the Eighth NTCIR Workshop

Noriko Kando
National Institute of Informatics, Tokyo JAPAN
kando@nii.ac.jp

ABSTRACT

For the Eighth NTCIR Workshop (NTCIR-8), we selected and organized seven research areas as "tasks" to investigate, test and benchmark newly constructed test collections. These areas are Complex Cross-Lingual Question Answering (CCLQA), Information Retrieval for Question Answering (IR4QA), Geographic and Temporal Search (GeoTime), Multilingual Opinion Analysis (MOAT), Patent Machine Translation (PATMT), Patent Mining (PATMN) and a pilot task for Community Question Answering (CQA). Among these, GeoTime and CQA are new to NTCIR and the others have some new elements. Each task had its own challenge to tackle and both organizers and participating research groups worked hard towards completing them. This paper provides an overview of the Eighth NTCIR Workshop, which is the latest in a series of community-oriented efforts to enhance related research by providing infrastructure for evaluation and testing, as well as a forum for researchers. The paper briefly describes the background, tasks, participants, test collections and other resources available through NTCIR-8. It is intended to serve as an introduction to the research described in detail in the full Proceedings of the Eighth NTCIR Workshop

Categories and Subject Descriptors

H.3.3 [**Information Systems**]: Information Search and Retrieval—retrieval models, search process**. General Terms:** Experimentation, Performance, Measurement **Keywords:** IR Evaluation, Cross-lingual Information Retrieval; Complex Question Answering, Geographic and Temporal Search, Geographic Information Retrieval, Opinion Analysis, Patent Machine Translation, Patent Mining, Text Mining, Community Question Answering

1. INTRODUCTION

The NTCIR Workshop [1] is a series of evaluation workshops designed to enhance research in information access (IA) technologies, including information retrieval (IR), cross-lingual information retrieval (CLIR), question answering (QA), automatic text summarization, text mining and so on, by providing test collections and a forum for researchers. It has been done as a collaborative community-based effort to build research infrastructure for testing and evaluation.

The aims of the NTCIR project are to

1. Encourage research in information access technologies by providing research infrastructure for evaluation and testing
2. Provide a forum for researchers interested in exchanging research ideas in an informal atmosphere
3. Investigate methodologies to evaluate information access technologies.

By providing infrastructure for large-scale experiments and evaluations, NTCIR has encouraged research on the selected IA tasks and facilitated research and technology transfer. The importance of such infrastructure in IA research has been widely recognized. Fundamental text processing procedures for IA, such as indexing, include language-dependent procedures. The NTCIR project therefore started in late 1997 with an emphasis on, but was not limited to, Japanese or other East Asian languages, and its workshop series has attracted international participation. For the eighth workshop, one hundred and eight research groups from eighty-five different organizations in seventeen different countries have registered to participate in experiments within one of the seven tasks. Among these, more than 75% are overseas participants.

In NTCIR, a unit of activity is one and a half years (i.e., eighteen months). The program committee selects several research areas called "tasks" from the task proposals submitted by research groups who wish to organize them. Each task is organized by these researchers of the domain as volunteer work and a task may consist of more than one subtask. Because we respect the interaction between participants, we consider the whole process – from discussion of the task design, to the call-for-task-participation, to the final meeting – as a "workshop".

From the previous workshop, NTCIR-7, we have introduced a "cluster structure" of tasks, where related tasks or those sharing a research target are grouped into a cluster, so that we can conduct a module-based evaluation as well as investigate the best combination of modules from different parties to search for a "dream system". NTCIR also encourages collaboration across tasks and clusters, and among research groups, for the purpose of more fruitful investigation.

1.1 Information Access

The term "information access" (IA) refers to the whole process from when a user recognizes his or her information needs, through the search for relevant documents, and then the application of information in relevant documents that are found. We have looked at IA technologies to help users apply information in large-scale document collections. IR, summarization and QA are part of a family of technologies aimed at the same target, although each has been investigated by rather different communities

1.2 Focus of NTCIR

From the beginning of the project, we have looked at both traditional laboratory-type IR system testing and the evaluation of challenging technologies. For the former, we have emphasized text retrieval and CLIR with Japanese or other East-Asian languages as well as testing on various document genres, and utilizing various types of parallel or comparable bilingual corpus. For the challenging issues, the target has been to shift from document retrieval to technologies that help users utilize

"information" in documents, and the investigation of methodologies and metrics for more realistic and reliable evaluation.

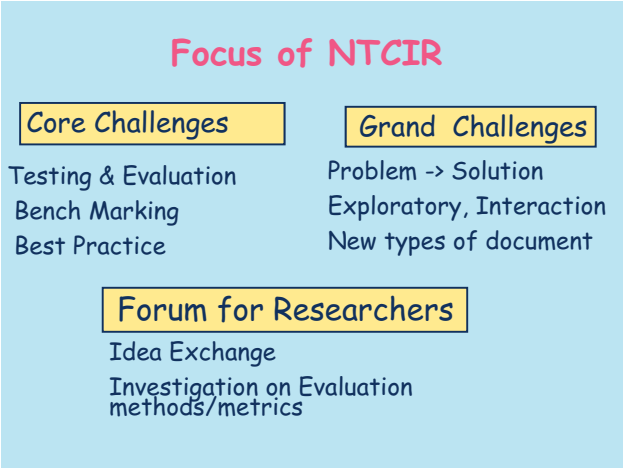

Figure 1. Focus of NTCIR Workshops

As shown in Figure 1, these two aspects can be rephrased as Core Challenges and Grand Challenges. Core challenges are where we want to encourage and leverage investigation by providing infrastructure for testing and evaluation, benchmarking, and showing the best practices for each technological challenge on the shared infrastructure. Grand challenges are where we believe collaborative research efforts aimed at a technological breakthrough can provide solutions for problems that we currently face in the real world. These two directions have been supported by a forum of researchers who are interested in cross-system comparison and collaboration.

For the core challenges, technologies and social environments are kept improving and enhancing, and then an appropriate research infrastructure for these new technologies must be continuously built and maintained by the research community. NTCIR thus has to continue the activity as a community-based collaborative effort.

For the grand challenges, the process may include a stage to explore the methodologies and the milestones and/or modules. From past NTCIRs, a series of patent mining/classification tasks targeting automatic construction of patent maps is an example of aiming in this direction. Another example is a series of multi-modal summarization of trend (MuST) tasks, where the aim is to encourage collaborative efforts to investigate and evaluate technologies for extracting trend-related information from text and summarizing and visualizing the information to encourage interactive exploratory information access by bridging different technological domains of IR, information extraction, visualization, and so on.

Grand challenges have included the investigation of and the research infrastructure to evaluate and test 1) interactive exploratory information access technologies, 2) information access for newly emerged document genres where the user's information needs, situation, usage and success criteria will differ from those for traditional information media, and 3) other domains where users have serious requirements for better information access and there is substantial space for investigation to improve the technological environment to support such serious information seeking and use.

The importance of the research on evaluation methodologies has been tremendously increased in these days. This trends partially resulted in many workshops on the topics were organized, papers on evaluation have been published at international major conferences and journals. When NTCIR was started, a large-scale document collection usable for research itself was valuable. In these days however it may not be so much difficult to obtain a large-scale dataset used in a specific research. And information access technologies has been matured and applied to wide variety of unique applications and tasks, and the need to test and evaluate such individual applications in individual environments are increased. Using crowd-sourcing assessments or aggregation of users judgments rather than creating one set of the exact right-answers has been becoming more and more appropriate to the information seeking tasks and contexts in the real-world.

Therefore, not only providing wide-variety of evaluation infrastructure, but also investigating the evaluation methodologies and then providing and suggesting the best practice for evaluation method and test collection building in particular individual setting have been more and more important as a activity of an evaluation workshop like NTCIR. For NTCIR, we have pursuit this direction by the organizers of each task, by encouraging the task participants to join such investigation by releasing submitted raw runs for research purposes, and 3) organizing a pre-meeting workshop called International Workshop on Evaluating Information Access (EVIA) on DAY-1 of each NTCIR meeting. TASKS AT NTCIR-8

2. OVERVIEW OF NTCIR-8 TASKS

For the Eighth NTCIR Workshop (NTCIR-8) [2-3], the process started with the call-for-task- participation in May 2009 and the meeting will be held on June 15-18, 2010 [4] at the National Institute of Informatics (NII) in Tokyo.

As shown in Figure 2, NTCIR-8 has selected seven "tasks" grouped into four clusters:

Cluster 1a: Advanced Cross-Lingual Information Access (ACLIA)
+ Complex Cross-Lingual Question Answering (CCLQA) task [5]
+ Information Retrieval for Question Answering (IR4QA) task [6]

Cluster 1b: Geographic and Temporal Search
+ Geographic and Temporal Search (GeoTime) task [New] [8]

Cluster 2: User Generated Contents
+ Multilingual Opinion Analysis Task (MOAT) [9]
+ [Pilot] Community Question Answering (CQA) [New] [10-11]

Cluster 3: Focused Domain : Patent
+ Patent Mining (PAT MN) task [12]
+ Patent Translation (PAT MT) task [13]
  + Evaluation subtask [New]

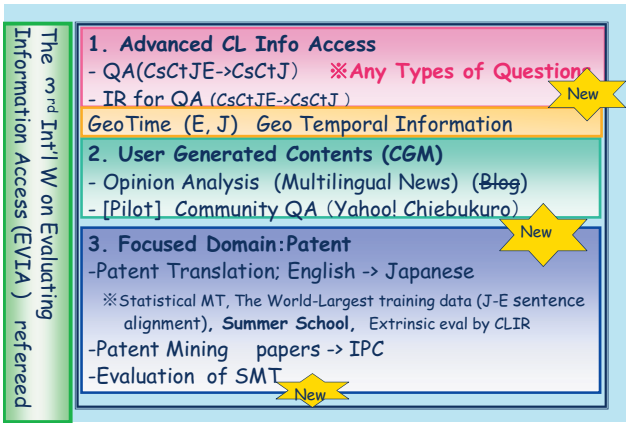

Figure 2. NTCIR-8 Task Clusters

Geographic and Temporal Search (GeoTime) is a new challenge, which is quite prevalent in various search and retrieval systems, including applications for the mobiles or search and visualization functionalities for digital libraries. There is potential usage and value for interactive user interfaces like interactive maps, which allow users not only to specify and narrow a search by geographic and temporal aspects, but also to explore and provide a view and note interesting events or trends. Although Geographic IR has been evaluated in CLE,F and there are several workshops and a special issue of an international journal on it, this is the first attempt to evaluate on cross-lingual access between an Asian language and English at NTCIR. There is also a potential extension toward the collaboration with ACLIA to investigate the place- or time-related questions, MOAT to investigate the differences and distribution of opinion regarding the same event or issue across different languages, and MuST to extract and visualize the relationship and tendencies over geographic and temporal facets.

Community QA is also a new challenge based on archives of real-world questions and communication between askers and answerers.

At NTCIR-8, ACLIA extended its scope to cover "any types of question" and evaluated all the questions including factoid-type by the same framework for the evaluation of more complex questions.

MOAT is a continuation from NTCIR-6's Opinion Analysis (OPINION) Pilot Task. At NTCIR-8 MOAT, the cross-lingual subtask was newly added. The task of NTCIR-8 MOAT has designed to an application for opinion question answering.

ACLIA, GeoTime and MOAT were tested on a new news story corpus which include the documents published in the years from 2002 to 2005, which consists of the New York Times for English, Xinhua Chinese for Simplified Chinese, UDN for Traditional Chinese and Mainichi News for Japanese.

Clusters 1 and 3 at NTCIR-8 are based on the extension of tasks well investigated in past NTCIRs with the focus of each investigation significantly changed at NTCIR-7. Cluster 1, or ACLIA, at NTCIR-7 and -8 is a module-based evaluation combining Complex Cross-lingual QA and IR for QA and it is an extension of CLIR, Question Answering Challenge (QAC) and CLQA from NTCIR-3 to -6. Evaluating "any type of question" is a challenge and even factoid type questions were evaluated within the same framework.

Furthermore, Cluster 3 proposed new research tasks at NTCIR-7 based on the prolonged experiences of Patent Retrieval and Patent Classification Targeting Text Mining for Automatic Patent Map Creation at NTCIR-3 to -6. NTCIR-8 is the second cycle of each of them (Figure 3).

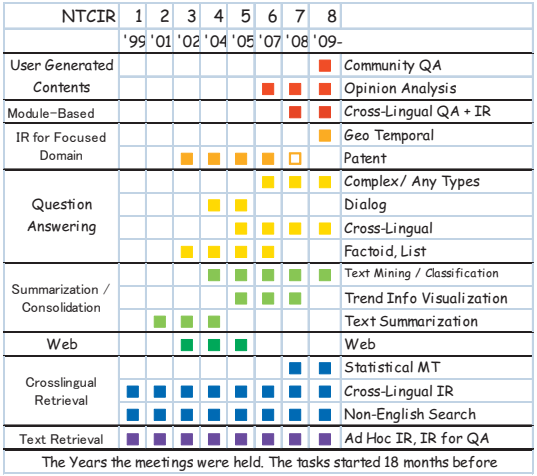

Figure 3. Tasks at NTCIR Workshop Series

2.1 Online Environment for assessment, annotation, and testing

For ACLIA, GeoTime, MOAT and CQA, online systems have been developed for relevance judgments and annotation, and these systems are available for collaborative efforts in distributed environments to build test collections and data sets which can be used for experiments. Such judgment and annotation works are best done in a region where the target language is used, so an online platform to enable such distributed work is essential.

With such platform, anybody can contribute the data creation. In CCLQA at ACLIA, dry run datasets were constructed by the participating research groups and some of the participants volunteered to contribute the judgments in formal runs as well. For GeoTime the task organizers and participating groups participated in relevance judgments. Traditional Chinese sub-tasks for CCLQA and IR4QA were supported by the task organizers from Taiwan.

Evaluation is critical for the researchers themselves. This is a kind of healthy cycle that the research community works collaboratively to build a research infrastructure of evaluation with some coordination. It is a natural direction as a community-based activities, I do thank for all the effort and contribution from the task organizers and participating groups for the community, and expect to see such efforts shall make NTCIR or any other evaluation campaign sustainable.

2.2 Module-based Evaluation

At NTCIR-7 and -8, the Advanced Cross-Lingual Information Access (ACLIA) cluster is a module-based evaluation combining Complex CLQA and IR for QA (Figure 4).

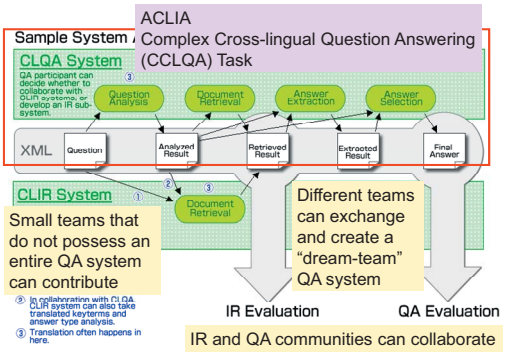

Figure 4. Module-based Evaluation Structure of ACLIA

The basic ideas behind the proposal of a module-based evaluation are to

- Develop advanced IA systems for any type of questions and enable the return of answers in relevant formats
- Enable glass-box evaluation. Researchers can easily learn from the experiments how and why each of modules contributed to the end-to-end system effectiveness and which module shall be improved..
- Build a "dream system" by bridging QA and IR communities
- Extend research infrastructure not only through traditional test collections, but also through API or modules
- Provide an opportunity for small research groups to make significant contributions

Not only the value as a module-based evaluation, each task in ACLIA has its own implication – in QA-viewpoint, evaluating any types of questions at CCLQA is a new challenge, and in IR viewpoint, IR4QA is a kind of focused retrieval for particular question types.

The framework was motivated from the detailed module-based analysis of the results of Cross-lingual Question Answering (CLQA) tasks at NTCIR-5 and -6. It revealed that one of the major problems in CLQA between Asian languages and English was the retrieval modules, especially the CLIR module, and the acute performance drops occurred in\\ them. CLQA systems generally worked well on question type analysis and answer extraction and formatting once the candidate documents were properly retrieved. However, in many cases, CLQA system failed to retrieve the appropriate document subsets that contained potential answers; this was especially the case for queries including proper names or any named entities which are often not included in ordinary translation dictionaries or machine translation systems.

On the other hand, CLIR has been investigated for a long time, from NTCIR-1 with shifting focuses until NTCIR-6. One of the major problems that all the participants have tackled is the solution of Out-of-Vocabulary (OOV) problems. In CLIR between languages using completely different alphabets, cognate matching cannot be used and searching for proper names is one of the most challenging tasks – proper names and named entities are often used in queries issued by users in daily life. Various methodologies for OOV have been proposed and tested through past NTCIRs.

If we can combine the technologies investigated in the QA community and the improved retrieval functionalities, including OOV solutions, tested in past CLIR tasks, we could conceivably realize a "dream system". Although CLQA, QAC and CLIR have been organized in the NTCIR series for a long time, the research communities for QA and IR are rather separate from each other, and few groups participate in both QA- and IR-oriented tasks. The organizers have proposed a module-based evaluation aimed at combining the best QA and the best IR systems. We also expect that such an environment will encourage participants to try other parts of the whole IA mechanisms.

CCLQA consists of highly complicated tasks made up of many modules. When we conduct an end-to-end evaluation, it is not easy for researchers to identify which parts are problematic in their highly complicated systems. If we can provide a module-based evaluation environment, it is much easier to identify the problems and to learn which parts of the systems need improvement and revisions. This module-based environment also encourages participation from small research groups or newcomers who can work on only part of the whole process of CLQA, but can work better on a focused module.

At NTCIR-7 ACLIA, we found good examples showing that the combination of IR and QA modules from different research groups produce better results than any other end-to-end CCLQA system. We expected to see more rich examples at NTCIR-8. Unfortunately, a number of participating groups could not submit their results in time for various reasons, so we could not see further examples of good combination runs at NTCIR-8. We hope the SEPIA framework that ACLIA has developed and provided for such module-based experimentation and evaluation and the test collections that have been developed will help and encourage the continuous effort to realize the benefits of such collaboration across different research groups. SEPIA will be released as an open source software.

ACLIA is one of the good examples of module-based evaluation, but we can easily find many other examples that have potentials to make a "dream system" and technological break-through. Information access technologies are highly complicated and consisted of various modules. In the technological domains investigated in NTCIR series, we can find other candidates. For example, statistical machine translation was especially worked well in the CLIR rather than human readings, and intellectual property experts showed strong expectation for CLIR than machine translation (MT) [14]. MT can contribute the better CLIR in many ways including query translation, incorporating

Table 1. Test collections constructed by NTCIR

Class	Collection	Task	Documents						Task data		
									Topic/ Question		Relevance
			Genre	Filename	Lang.	Year	# of docs	Size	Lang.	#	judge
CLIR on Sciemtific	NTCIR-1	IR	Sci. abstract	ntc1-je ++	JE	1988-1997	339,483	577MB	J	83	3 grades
				ntc1-j ++	J		332,918	312MB			
				ntc1-e ++	E		187,080	218MB		60	
		TE*5		ntc1-tmrc ++	J		2,000	-		-	-
	NTCIR-2	IR	Sci. abstract	ntc2-j ++	J	1986-1999**	400,248	600MB	JE	49	4 grades
				ntc2-e ++	E		134,978	200MB			
CLIR on News	CIRB010	IR	News	CIRB010	C_t	1998-1999	132,173	132MB	C_tE	50	4 grades
	NTCIR-3 CLIR	IR	News	KEIB010	K	1994	66,146	74MB	C_tKJE	30	4 grades
				CIRB011	C_t	1998-1999	132,173	870MB	C_tKJE	50	4 grades
				CIRB020 ++			249,508				
				Mainichi +&	J		220,078				
				EIRB010	E		10,204				
				Mainichi Daily ++			12,723				
	NTCIR-4 CLIR	IR	News	CIRB011	Ct	1998-1999	132,173	ca.3GB	CtKJE	60	4 grades
				CIRB020 ++			249,203				
				Hankookilbo ++	K		149,921				
				Chosenilbo ++			104,517				
				Mainichi +&	J		220,078				
				Yomiuri +&			373,558				
				EIRB010	E		10,204				
				Mainichi Daily ++			12,723				
				Korea Times ++			19,599				
				Hong Kong Standard ++			96,683				
				Xinhua +&			208,167				
	NTCIR-5 CLIR	IR	News	CIRB040r ++	Ct	2000-2001	901,446		CtKJE	50	4 grades
				Hankookilbo ++	K		85,250				
				Chosenilbo ++			135,124				
				Mainichi +&	J		199,681				
				Yomiuri +&			658,719				
				Mainichi Daily ++	E		12,155				
				Korea Times ++			30,530				
				Daily Yomiuri +&			17,741				
				Xinhua +&			198,624				
	NTCIR-6 CLIR	IR	News	CIRB040r ++	Ct	2000-2001	901,446		CtKJE	50 (selected from NTCIR-3,4)	4 grades
				Hankookilbo ++	K		85,250				
				Chosenilbo ++			135,124				
				Mainichi +&	J		199,681				
				Yomiuri +&			658,719				
ACLIA	NTCIR-7 ACLIA	IR+QA	News	CIRB020 ++	Ct	1998-1999	249,508		CJE		3 grades
				CIRB040r ++	Ct	2000-2001	901,446				
				Lianhe Zaobao ++	Cs	1998-2001	249,287				
				Xinhua +&	Cs	1998-2001	295,875				
				Mainichi +&	J	1998-2001	419,759				
	NTCIR-8 ACLIA	IR+QA	News	Xinhua +&	Cs	2002-2005	308,845		CsCtJE	100	3 grades
				UDN ++	Ct	2002-2005	1,663,517				
				Mainichi +&	J	2002-2005	377,941				
CLQA	NTCIR-5 CLQA	QA	News	CIRB040r ++	C	2000-2001	901,446		CJE	smpl:300, test:200 *6	3 grades *7
				Yomiuri +&	J		658,719				
				Daily Yomiuri +&	E		17,741				
	NTCIR-6 CLQA	QA	News	CIRB020 ++	Ct	1998-1999	249,203		CJE	J-E/J-J/E-J: 200, C-E/C-C/E-C/E-E: 150	3 grades *7
				Mainichi +&	J		220,078				
				EIRB010	E		10,204				
				Mainichi Daily ++			12,723				
				Korea Times ++			19,599				
				Hong Kong Standard ++			96,683				
CQA	NTCIR-8 CQA	QA	QA site on Web	Yahoo! Q&A corpus (Chiebukuro) ++	J	Apr.2004 to Oct.2005					
OPINION	NTCIR-6 OPINION	IE/analysis	News	CIRB020 ++	Ct	1998-2001	249,203	32, 30, 28 doc sets for C, J, E, respectively, were selected from those described as left	CtKJE	32 (selected from NTCIR-3,-4,-5 CLIR)	2 types, 3 metrics
				CIRB040r ++			901,446				
				Mainichi +&	J		419,759				
				Yomiuri +&			1,032,277				
				EIRB010	E		10,204				
				Mainichi Daily ++			242,878				
				Korea Times ++			50,129				
				Hong Kong Standard ++			96,683				
				Xinhua +&			409,971				
	NTCIR-7 MOAT	IE/analysis	News	CIRB020 ++	Ct	1998-1999	249,508		CtCsJE		2 types, 3 metrics
				CIRB040r ++	Ct	2000-2001	901,446				
				Lianhe Zaobao ++	Cs	1998-2001	249,287				
				Xinhua +&	Cs		295,875				
				Mainichi +&	J		419,759				
				Mainichi Daily ++	E		24,878				
				Korea Times ++			50,129				
				Hong Kong Standard ++		1998-1999	96,683				
				Straits Times++		1998-2001	-				
				Xinhua +&		1998-2001	406,791				
	NTCIR-8 MOAT	IE/analysis	News	Mainichi +&	J	2002-2005			CtCsJE		
				New York Times +&	E	2002-2005	315,417				
				UDN2002-2005 (United Daily News, United Express, Ming Hseng News, Economic Daily News, Star News) ++	Ct	2002-2005	1663517				
				Xinhua Chinese +&	Cs	2002-2005	308,845				

Table 1. Test collections constructed by NTCIR (continued)

Class	Collection	Task	Documents						Task data		
									Topic/ Question		Relevance
			Genre	Filename	Lang.	Year	# of docs	Size	Lang.	#	judge
GeoTime	NTCIR-8 GeoTime	IE/analysis	News	Mainichi +&	J	2002-2005					
				New York Times +&	E	2002-2005	315,417				
MuST	NTCIR-6 MuST	IE/analysis	News	Mainichi +&	J	1998-1999	220,078	260MB			
	NTCIR-7 MuST	IE/analysis	News	Mainichi +&	J	1998-2001	419,759	535MB			
Patent	NTCIR-3 PATENT	IR	Patent full	kkh *3 ++	J	1998-1999	697,262	18GB	C_tC_sK JE	31	3 grades
			Abstract	jsh *3 ++	J	1995-1999	1,706,154	1,883MB			
			Abstract	paj *3 ++	E	1995-1999	1,701,339	2,711MB			
	NTCIR-4 PATENT	IR	patent fulltext	Publication of unexamined patent application ++	J	1993-1997	ca. 1,700,000	ca.27GB	E	Main: 34, Add: 69	3 grades
			patent abstract	Patent Abstracts of Japan (PAJ) ++	E	1993-1997	ca. 1,700,000	ca.5GB			
	NTCIR-5 PATENT	IR	patent fulltext	Publication of unexamined patent application++	J	1993-2002	3,496,252	ca.45GB	JE	34+1189 in NTCIR-5, added 349+1681in	3 grades
			patent abstract	Patent Abstracts of Japan (PAJ) ++	E	1993-2002	3,496,252	ca.10GB			
	NTCIR-6 PATENT	IR	patent fulltext	Patent grant data published from USPTO++	E	1993-2000	981,948		E	3221	3 grades
Patent Mining	NTCIR-7 PATMN	Mining	patent fulltext	Publication of unexamined patent application++	J	1993-2002	3,496,252	94.5GB	JE	Japanese/ Cross-lingual (E2J) 976	2
				Patent Abstracts of Japan(PAJ)++	E		3,496,252	ca.5GB			
				Patent grant data published from USPTO++	E		1,315,470	52.6GB			
			Scientific abstract	ntc1-je++	JE	1998-1997	339,483	577MB		English/ Cross-lingual (J2E) 976	2
				ntc1-j++	J		332,918	312MB			
				ntc1-e++	E		187,080	218MB			
			Scientific abstract	ntc2-j++	J	1986-1999	400,248	600MB			
				ntc2-e++	E		134,978	200MB			
	NTCIR-8 PATMN	Mining	patent fulltext	Publication of unexamined patent applications++	J	1993-2002	3,496,252				
				Patent Abstracts of Japan(PAJ)++	E		3,496,252				
				Patent grant data published from USPTO++	E		981,948				
			Scientific abstract	ntc1-je++	JE	1998-1997	339,483	577MB			
				ntc1-j++	J		332,918	312MB			
				ntc1-e++	E		187,080	218MB			
			Scientific abstract	ntc2-j++	J	1986-1999	400,248	600MB			
				ntc2-e++	E		134,978	200MB			
Patent Translation	NTCIR-7 PATMT	MT	patent fulltext	Publication of unexamined patent application++	J	1993-2002	3,496,252	94.5GB	JE		
				Patent grant data published from USPTO++	E	1993-2002	1,315,470	54.6GB			
	NTCIR-8 PATMT	MT	patent fulltext	Publication of unexamined patent applications++	J	1993-2007			JE		
				Patent grant data published from USPTO++	E	1993-2007					
QA	NTCIR-3 QA	QA	News	Mainichi +&	J	1998-1999	220,078	260MB	J*	1200	exact answer
	NTCIR-4 QA	QA	News	Mainichi +&	J	1998-1999	220,078	ca.776MB	J*	197	exact answer
				Yomiuri +&			373,558			199	
										251	
	NTCIR-5 QA	QA	News	Mainichi +&	J	2000-2001	199,681	260MB	J*	50 series (360 Q)	graded
	NTCIR-6 QA	QA	News	Mainichi +&	J	1998-2001	419,759	535MB	J	100 Q (any kind of Q)	graded (3 types, 4 levels)
	NTCIR-7 ACLIA(CCLQA)	QA	News	CIRB020 ++	Ct	1998-1999	249,508	320MB	CJE	100 Q (any kind of Q)	
				CIRB040r ++	Ct	2000-2001	901,446	582MB			
				Lianhe Zaobao ++	Cs	1998-2001	249,287	411MB			
WEB	NTCIR-3 WEB	IR	Web (html/text)	NW100G-01++	multiple* 4	crawled in 2001	11,038,720	100GB	J*	47	4 grades + relative
				NW10G-01++			1,445,466	10GB			
	NTCIR-4 WEB	IR	Web (html/text)	NW100G-01++	multiple* 4	crawled in 2001	11,038,720	100GB	J*		3 grades
	NTCIR-5 WEB	IR	Web (html/text)	NW1000G-04++	multiple* 4	crawled in 2004	98,870,352	1.36TB	J*	269 + 847	3 grades
OTHERs	available for future task		QA site on Web	Yahoo! Q&A corpus (Chiebukuro) ++	J	Apr.2004 to Oct.2005					
			News	Singpore Press ++	Cs	1998-2001					

J:Japanese, E:English, C:Chinese (C_t:Traditional Chinese, C_s: Simplified Chinese), K:Korean;
All the topics/questions and relevance judgements/answers are available for research purpose for free
"++" indicates the document collections available from NII for research purpose
"+&" indicates the document collections available for task participants for free, and available for research purpose use from research purpose other than NTCIR participation from other party with fee
* English translation is available
** gakkai subfiles: 1997-1999, kaken subfiles: 1986-1997
*3: kkh : Publication of unexamined patent application, jsh: Japanese abstract, paj: English translation of jsh
*4: almost Japanese or English (some in other languages)
*5: Term extraction/ role analysis
*6: 300+200 questions for C documents, and 300+200 questions for JE documents
*7: Right, Unsupported, Wrong

translation probability to IR weight, bi-directoral approach both query and documents are translated, translation of retrieved documents for relevance judgments in interactive and exploratory "human-in-the-loop" environment of real world patent retrieval. How to combine translation and IR components towards the best effectiveness in CLIR has not been fully investigated.

Another example is MuST – Multi-modal Summarization of Trends which was conducted at NTCIR6-7, where information extraction of the trend-related information from the document collections and visualization modules which belong to completely different technical domains are essential.

3. Test Collections

Test collections and data set built in the past NTCIRs are listed in Table 1. Most of them are available for research purpose for non-participants. The test collections constructed through NTCIR-8 will be available for research purpose for outside the NTCIR's task participants after the NTCIR-8 Meeting although some of the news document collections will have to be obtained the license from different providers

4. Participants

Table 2 is a list of the active participating research groups in the NTCIR-8. One hundred and eight groups from eighty-five organization registered, and sixty-five from seventeen different countries and areas were remained as active participants up to the final meeting.

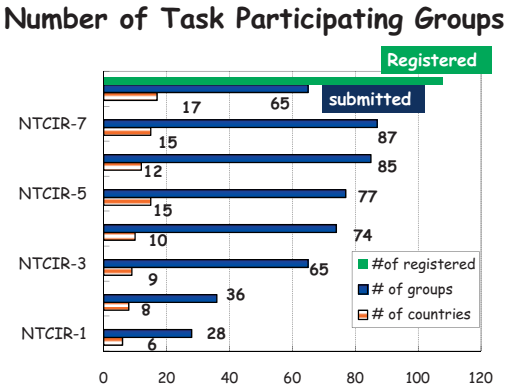

Figure 5. Number of Participating Groups

As shown in Figures 5 the numbers of participants who completed tasks declined at NTCIR-8 and we are now investigating why many of the registered groups could not submitted the results. We hope to use our findings to improve the organization in the future NTCIRs. Different tasks attracted different research groups. Many international participants enrolled in ACLIA, MOST, and PATMN. GeoTime attracted many new comers joining NTCIR.

[CCLQA]
Carnegie Mellon Univ
Dalian Univ of Technology
National Taiwan Ocean Univ
Shenyan Institute of Aeronautical Engineering
Univ of Tokushima
Wuhan Univ

[IR4QA]
Carnegie Mellon Univ
Chaoyang Univ of Technology
Dalian Univ of Technology
Dublin City Univ
Inner Mongolia Univ
Queensland Univ of Technology
Shenyan Inst of Aeronautical Engineering
Trinity College Dublin
Univ California, Berkeley
Wuhan Univ
Wuhan Univ (Computer School)
Wuhan Univ of Science and Technology

[GeoTime]
Dublin City Univ
Hokkaido Univ
INESC-ID, Porugal
International Inst of Technology, Hyderbad
Kieo Univ
Nataional Inst of Materials Science
Osaka Kyoiku Univ
Univ California, Berkeley
Univ of Iowa
Univ of Lisbon
Yokohama City Univ

[MOAT]
Beijing Uni of Posts and Telecommunications
Chaoyang Univ of Technology
Chinese Univ of HK+ Tsinghua Univ
City Univ of Hong Kong (2 groups)
Hong Kong Polytechnic Univ
KAIST
National Taiwan Univ
NEC Laboratories China
Peking Univ
Pohang Univ of Sci and Tech
SICS
Toyohashi Univ of Technology
Univ of Alicante
Univ of Neuchatel
Yuan Ze Univ

[Patent Mining]
Hiroshima City Univ
Hitachi, Ltd.
IBM Japan, Ltd.
Institute of Scientific and Technical Information of China
KAIST
National Univ of Singapore
NEC
Shanghai Jiao Tong Univ
Shenyang Institute of Aeronautical Engineering
Toyohashi Univ of Technology
Univ of Applied Sciences - UNIGE

[Patent Translation]
Dublin City University, CNGL
Hiroshima City University
Kyoto University
NiCT
Pohang Univ of Sci and Tech
tottori university
Toyohashi University of Technology
Yamanashi Eiwa College

[Community QA]
Microsoft Research Asia
National Institute of Informatics
Shirayuri College

Table 2. Active Participating Groups of the Seventh NTCIR Workshop

5. Summary

This has been a brief overview of the Eight NTCIR Workshop. The details of the achievements from each task and those of each participant are presented at the NTCIR-8 Meeting and reported in the collected papers in the proceedings CR and the online proceedings [3]. At the time of NTCIR-8 Meeting, both of CD and online version of the proceedings are identical, but the online proceedings will be richer by adding enhanced versions and the slides and posters.

The test collections used in the tasks of the NTCIR-8 and the archives of the system-produced submitted runs will be available for research purpose. We expect that many of the research groups involved in the larger NTCIR community will work collaboratively to investigate the system mechanisms and to analyze further results, consequently learning from each other's experience.

Although WEB was not considered at NTCIR-8, this was not because of a lack of interest. In fact, we are interested in various aspects of evaluating information access on WEB. WEB has changed incredibly in many ways, both in the quality and the quantity of the information access in the real world, and the technologies to support access. As a wide variety of users in this information world search, learn and investigate on the Web and other information resources, interactive and exploratory information access and its evaluation is natural extension of our efforts. We then had a break for better planning for the future tasks investigating various aspects of the technologies to support interactive and exploratory search on the Web as well as those with other traditional or newly emerged document genres for various usage and information seeking tasks behind.

Evaluation must adapt to technological evolution and changes in social needs. We are working towards this goal together with the wider research community through a collaborative community-based effort. Any leads and suggestions are always welcome.

6. References

[1] NTCIR Project: http://research.nii.ac.jp/ntcir/

[2] NTCIR Workshop 8 (2009-2010) NTCIR-8 Meeting: The 8th NTCIR Workshop Meeting on Evaluation of Information Access Technologies: Information Retrieval, Question Answering and Cross-Lingual Information Access, June 15-18, 2010. http://research.nii.ac.jp/ntcir/ntcir-ws8/meeting/

[3] Daisuke Ishikawa, Noriko Kando, Kazuaki Kishida, Teruko Mitamura, Hidetsugu Nanba, Tetsuya Sakai, Yohei Seki, Miho Sugimoto, Takehiro Utsuro (eds): NTCIR Workshop 8: Proceedings of the Etighth NTCIR Workshop Meeting on Evaluation of Information Access Technologies: Information Retrieval, Question Answering and Cross-Lingual Information Access, Tokyo Japan, June 15-18, 2010, NII, Tokyo 2010. (ISBN: 978-4-86049-053-9) Available online: http://research.nii.ac.jp/ntcir/ workshop/OnlineProceedings8/

[4] NTCIR Workshop 8 Meeting (June 15-18, 2010) http://research.nii.ac.jp/ntcir/ntcir-ws8/meeting/

[5] Teruko Mitamura, Hideki Shima, Tetsuya Sakai, Noriko Kando, Tatsunori Mori, Kohichi Takeda, Chin-Yew Lin, Ruihua Song, Chuan-Jie Lin and Cheng-Wei Lee. Overview of the NTCIR-8 ACLIA Tasks: Advanced Cross-Lingual Information Access. In Proceedings of the Eighth NTCIR Workshop Meeting, Tokyo Japan, June 15-18, 2010, NII, Tokyo 2010, pp. 15-24.

[6] Tetsuya Sakai, Hideki Shima, Noriko Kando, Ruihua Song, Chuan-Jie Lin, Teruko Mitamura, Miho Sugimoto, Cheng-Wei Lee. Overview of the NTCIR-8 ACLIA IR4QA. In Proceedings of the Eighth NTCIR Workshop Meeting, Tokyo Japan, June 15-18, 2010, NII, Tokyo 2010, pp. 63-93.

[7] Hideki Shima. Teruko Mitamura. SEPIA (Standard Evaluation Package for Information Access). Presented at NTCIR-8 Meeting Poster Session ACLIA-01. June 15-18, 2010.

[8] Fredric Gey, Ray Larson, Noriko Kando, Jorge Machado and Tetsuya Sakai. NTCIR-GeoTime Overview: Evaluating Geographic and Temporal Search, In Proceedings of the Eighth NTCIR Workshop Meeting, Tokyo Japan, June 15-18, 2010, NII, Tokyo 2010, pp. 147-153.

[9] Yohei Seki, Lun-Wei Ku, Le Sun, Hsin-Hsi Chen and Noriko Kando. Overview of Multilingual Opinion Analysis Task at NTCIR-8. In Proceedings of the Eighth NTCIR Workshop Meeting, Tokyo Japan, June 15-18, 2010, NII, Tokyo 2010, pp. 209-220.

[10] Daisuke Ishikawa, Tetsuya Sakai and Noriko Kando. Overview of the NTCIR-8 Community QA Pilot Task(Part I). In Proceedings of the Eighth NTCIR Workshop Meeting, Tokyo Japan, June 15-18, 2010, NII, Tokyo 2010, pp. 421-432.

[11] Tetsuya Sakai, Daisuke Ishikawa and Noriko Kando. Overview of the NTCIR-8 Community QA Pilot Task(Part II). In Proceedings of the Eighth NTCIR Workshop Meeting, Tokyo Japan, June 15-18, 2010, NII, Tokyo 2010, pp. 433-457.

[12] Atsushi Fujii, Masao Utiyama, Mikio Yamamoto, Takehito Utsuro, Terumasa Ehara, Hiroshi. Overview of the Patent Translation Task at the NTCIR-8 Workshop. In Proceedings of the Eighth NTCIR Workshop Meeting, Tokyo Japan, June 15-18, 2010, NII, Tokyo 2010, pp. 371-376.

[13] Hidetsugu Nanba, Atsushi Fujii, Makoto Iwayama and Taiichi Hashimoto. Overview of the Patent Mining Task at the NTCIR-78 Workshop. In Proceedings of the Eighth NTCIR Workshop Meeting, Tokyo Japan, June 15-18, 2010, NII, Tokyo 2010, pp. 293-302.

[14] The Second Information Retrieval Facility Symposium, Vienna, Nov. 2008.