

行政院國家科學委員會專題研究計畫 成果報告

跨語言文字資訊系統之量化評估(II)

計畫類別：個別型計畫

計畫編號：NSC92-2213-E-002-077-

執行期間：92年08月01日至93年07月31日

執行單位：國立臺灣大學圖書資訊學系暨研究所

計畫主持人：陳光華

報告類型：精簡報告

處理方式：本計畫可公開查詢

中 華 民 國 93 年 8 月 23 日

行政院國家科學委員會專題研究計畫報告

跨語言文字資訊系統之量化評估(II)

計畫編號：NSC92-2213-E-002-077

執行期限：92年8月1日至93年7月31日

主持人：陳光華

國立臺灣大學圖書資訊學系 副教授

1. Chinese Abstract

透過系統評估的程序，研究者能藉以驗證跨語言資訊檢索系統的效益、比較各種檢索技術的優劣，以作為改進之參考，使資訊檢索系統的運作及效能更臻完善。本計畫為國際性合作研究計畫，目標為中、日、韓、英等多語言資訊檢索系統的評估，參與國家有日本、韓國、台灣，分別由各國的學術研究機構支援。在計畫執行過程，各國皆需要蒐集各項文件資源，建構使用者的資訊需求，並建置共同的評估平台，同時交換所蒐集的資源並進行技術交流。今年為本計畫的第二年，第一年計畫已與日、韓兩國共同舉辦一次跨語言資訊檢索系統的評估會議，共有來自8個國家地區23個研究團隊參與該項評估會議；第二年計畫亦舉辦了評估會議，共有來自9個國家地區26個研究團隊參與評估會議。

2. Abstract

Most of researchers focus on the design of new CLIR systems using complicated technologies rather than the evaluation of the designed systems. However, evaluation is indispensable for a successful information system. The project is an international joint effort of Japan, Korea, and Taiwan. The Japanese partner will be supported by National Institute of Informatics; the Korean partner will find the funding support from Korean Science and Engineering Foundation; the Taiwan partner is supported by National Science Council. Each country will collect the resources for her language independently, construct the common evaluation platform cooperatively, and exchange the resources and technologies. This year is the second year of the project. One evaluation forum has been held in the first-year project. Totally, 23 research groups from 8 countries participates this forum. The second year project also holds an evaluation forum. Totally 26 groups from 9 countries participates this forum.

3. Introduction

Cross-lingual information retrieval (CLIR) is recognized as an important research issue for the information society in which the Internet spreading globally in the world plays a crucial role. In order to promote moreover research efforts on CLIR among a set of East-Asian (i.e., Chinese, Japanese and Korean) and English languages, the CLIR task is again organized in the NTCIR-4 project after the NTCIR-3 CLIR task was completed on October 2002.

The NTCIR-4 CLIR task has taken over three subtasks from the previous task, i.e.,

- *Multilingual CLIR* (MLIR),
- *Bilingual CLIR* (BLIR), and
- *Single Language IR* (SLIR).

Furthermore, in the NTCIR-4 CLIR task, we have a new subtask, *Pivot Bilingual CLIR* (PLIR). PLIR is a special kind of BLIR in which a third language is employed as an intermediary for translating the source language into the target one, e.g., for Chinese-Japanese BLIR, Chinese query terms are translated into English words and then the set of English words is translated into Japanese words. The PLIR task is set to contribute toward solving problems of insufficient language resources for directly translation between East-Asian languages.

The document sets are extended for the NTCIR-4 CLIR task, i.e., 1998-99 Korean document records are added to the test collection (in the previous NTCIR-3, while Chinese (C), Japanese (J) and English (E) documents published in 1998 and 1999 are available, the publishing year of Korean (K) documents was 1994). The addition of documents allows us to conduct experiments on a CJKE multilingual collection. Also, Japanese and English document sets are also augmented so that the size of the document sets in each language is well balanced (In the NTCIR-3 collection, the English part is relatively small).

4. Design of the CLIR Task

4.1 Multilingual CLIR (MLIR). In general, the document set of MLIR subtask consists of two or more languages. For the NTCIR-4 CLIR task, the participants are allowed to submit results of runs for two types of multilingual document collection,

- "*Large collection*": CJKE collection, which consists of Chinese(C), Japanese(J), Korean(K) and English(E) documents
- "*Small collection*": CJE collection, which consists of Chinese(C), Japanese (J), and English (E) documents.

Regarding the topic set, participants can select one language from CJKE for each run. Therefore, there are eight combinations of topic sets and document sets, i.e., Topic set: C or J or K or E >> Doc set: CJKE or CJE.

4.2 Bilingual CLIR (BLIR). BLIR means that the document set in a single language is searched for a topic in a different language, e.g., searching Japanese documents for Korean topics (K -> J run). In the NTCIR-4 CLIR task, participants are basically not allowed to submit results of runs using topics written

in English, except the case of trying pivot language approach (i.e., PLIR).

4.3 Pivot Bilingual CLIR (PLIR). This subtask is a new challenge at the NTCIR-4 CLIR task. As already mentioned, this approach employs a third language as an intermediary for translation of query or document texts. Also, the participants submitting runs for this subtask are allowed to also submit BLIR runs using English topics (i.e., E -> C or J or K) in order to analyze comparatively performance of the approach.

4.4 Single Language IR (SLIR). The topic set and document sets of SLIR are written in a same language.

4.5 Types of runs. Basically, each topic consists of four fields, i.e., “T” (TITLE), “D” (DESC), “N” (NARR) and “C” (CONC) (see below for details). We can categorize search runs based on the fields used for execution. In the NTCIR-4 CLIR task, the following types of runs are adopted:

- *Mandatory runs:* T-run and D-run
- *Recommended runs:* DN-run
- *Optional runs*

Other any combinations of fields are allowed to submit as optional runs according to each participant’s research interests, e.g., TDN-run, DC-run, TDNC-run and so on.

5. Test Collection

5.1 Document Sets

The documents are news articles collected from various news agencies from different countries. Table 1 shows the sources and numbers of records in the document collections. The tags used for separating each field are also indicated in Table 2.

Table 1 Document sets

Sources	No. of Docs	
<i>Chinese 1998-99</i>		
CIRB020 (United Daily News)	249,508	
CIRB011 (China Times, China Times Express, Commercial Times, China Daily News, Central and Daily News)	132,173	
Total	381,681	
<i>Japanese 1998-99</i>		
Mainichi	220,078	
Yomiuri	375,980	
Total	596,058	
<i>Korean 1998-99</i>		
Hankookilbo	149,921	
Chosunilbo	104,517	
Total	254,438	
<i>English 1998-99</i>		
EIRB010	Taiwan News	7,489
	China Times English News (Taiwan)	2,715
Mainichi Daily News (Japan)	12,723	
Korea Times	19,599	
Xinhua (AQUAINT)	208,168	
Hong Kong Standard	96,856	
Total	347,550	

Table 2 Tags Set for Documents

<i>Mandatory tags</i>	
<DOC>	The tag for each document
<DOCNO>	Document identifier
<LANG>	Language code: CH, EN, JA, KR
<HEADLINE>	Title of this news article
<DATE>	Issue date
<TEXT>	Text of news article
<i>Optional tags</i>	
<P>	Paragraph marker
<SECTION>	Section identifier in original newspapers
<AE>	Contain figures or not
<WORDS>	Number of words in 2 bytes (for Mainichi Newspaper)

5.2 Topics

The following shows a sample topic.

<TOPIC>
<NUM>009</NUM>
<SLANG>CH</SLANG>
<TLANG>EN</TLANG>
<TITLE>Japan, South Korea, Fishery Agreement</TITLE>
<DESC>Find articles on the content of the final fishery agreement between Japan and South Korea</DESC>
<NARR>
<BACK>There are frequent disputes between Japan and South Korea because of the 35 years of colonized reign. Things worsened in January of 1998 when Japan announced the abolishment of the fishery agreement of 1965. Finally, in September of 1998, a new fishery agreement between Japan and South Korea was reached despite disputes over the sovereignty of the isles. It marked an end to eight months of serious disputes between the two countries. Please query the content of this new agreement for things such as allocation of fishing areas and results of negotiation.</BACK>
<REL>Documents of reports on the final fishery agreement are relevant. Reports on historical disputes and events between Japan and South Korea are not relevant.</REL>
</NARR>
<CONC>Japan, South Korea, Fishery Agreement, Isles, Fishing Area</CONC>
</TOPIC>

The tags used in topics are shown in Table 3. The topics were created in Taiwan, Japan and Korea, separately, and finally 60 topics were selected based on results of feasibility test checking the numbers of relevant documents in each document set. The original language used in the process of creating topics is recorded in the <SLANG> field.

Table 3 Tag Set for Topics

<TOPIC>	The tag for each topic
<NUM>	Topic identifier
<SLANG>	Source language code: CH, EN, JA, KR
<TLANG>	Target language code: CH, EN, JA, KR
<TITLE>	The concise representation of information request, which is composed of noun or noun phrase.
<DESC>	A short description of the topic. The brief description of information need, which is composed of one or two sentences.
<NARR>	The <NARR> may have three parts; (1)<BACK>...</BACK>: background information about the topic is described. (2)<REL>...</REL>: further interpretation of the request and proper nouns, the list of relevant or irrelevant items, the specific requirements or limitations of relevant documents are given. (3)<TERM>...</TERM>: definition or explanation of proper nouns, scientific terms and so on.
<CONC>	The keywords relevant to whole topic.

Subsequently, selected 60 topics were translated into English, and each English topic was translated into each Asian language except the original language. All translation works were done by human translators. Through the process, four languages (CJKE) versions of all 60 topics were prepared.

6 Evaluations

6.1 Submission of Results

In total, search results were submitted by 26 groups from 9 countries. Regarding the numbers of participants, Japan is dominant (9 groups), followed by USA (6 groups), China (3 groups), and Taiwan (2 groups).

Unfortunately, 7 groups that applied to participate in the NTCIR-4 CLIR task could not submit final results for some reasons. In total, 368 runs were submitted, of which 182 (49.5%) are for SLIR, 149 (40.5%) are for BLIR (including PLIR), and 37 (10.1%) are for MLIR.

6.2 Effective sets of topics for evaluation

There are some topics for which relevant documents are very few. Therefore, the task organizers were determined to employ again so-called "3-in-S+A" criterion, which was applied at the NTCIR-3. The "3-in-S+A" criterion means that only the topics having three or more "rigid" relevant documents are used for evaluation.

According to this criterion, the sets of topics for each document collection are as follows.

1. Topics for SLIR BLIR and PLIR
 - (1) Chinese Collection (C): 59 topics (ID: 001-024, 026-060) are used for evaluation. The topic 025 is removed.
 - (2) Japanese Collection (J): 55 topics (ID 003-021, 023-024, 026-037, 039-060) are used for evaluation. The topics, 001, 002, 022, 025 and 038, are removed.
 - (3) Korean Collection (K): 57 topics (ID: 002-009, 012-060) are used for evaluation. The topics, 001, 010 and 011, are removed.
 - (4) English Collection (E): 58 topics (ID: 002-037, 039-060) are used for evaluation. The topics, 001 and 038, are removed.
2. Topics for MLIR
CJE and CJKE: All 60 topics are used (no topic is removed).

7. Scoring for Search Results

7.1 SLIR runs

7.1.1 C-C runs. In total, 52 C-C monolingual runs were submitted by 13 groups. Table 4 shows average, median, maximum and minimum values of mean average precision (MAP) by types of runs. We use the following notations;

- C-C: all C-C monolingual runs
- C-C-T: all C-C <TITLE>-only runs (T-runs)
- C-C-D: all C-C <DESC>-only runs (D-runs)
- C-C-O: all other runs than T- or D-runs

I2R-C-C-D-01 based on ontological query expansion is dominant. It seems that there are almost no statistically significant differences between the other seven groups.

7.1.2 J-J runs. In total, 58 J-J monolingual runs were submitted by 14 groups. Table 5 shows average, median, maximum and minimum values of mean average precision (MAP) by types of runs.

7.1.3 K-K runs. In total, 31 K-K monolingual runs were submitted by 8 groups. Table 6 shows average, median, maximum and minimum values of mean average precision (MAP) by types of runs.

7.1.4 E-E runs. In total, 41 E-E monolingual runs were submitted by 10 groups. Table 7 shows average, median, maximum and minimum values of mean average precision (MAP) by types of runs.

Table 4 MAP of overall C-C runs

(a) Average and median

	Average		Median	
	Rigid	Relax	Rigid	Relax
C-C	0.1985	0.2471	0.1999	0.2537
C-C-T	0.1943	0.2378	0.1881	0.2356
C-C-D	0.1826	0.2328	0.1741	0.2219
C-C-O	0.2230	0.2762	0.2363	0.2915

(b) Min and max

	Min		Max	
	Rigid	Relax	Rigid	Relax
C-C	0.1251	0.1548	0.3255	0.3880
C-C-T	0.1327	0.1638	0.3146	0.3799
C-C-D	0.1251	0.1548	0.3255	0.3880
C-C-O	0.1461	0.1774	0.2556	0.3103

Table 5 MAP of overall J-J runs

(a) Average and median

	Average		Median	
	Rigid	Relax	Rigid	Relax
J-J	0.3258	0.4247	0.3358	0.4329
J-J-T	0.3114	0.4073	0.3135	0.4112
J-J-D	0.3227	0.4212	0.3352	0.4295
J-J-O	0.3441	0.4467	0.3487	0.4622

(b) Min and max

	Min		Max	
	Rigid	Relax	Rigid	Relax
J-J	0.1966	0.2759	0.3915	0.4963
J-J-T	0.1966	0.2759	0.3890	0.4864
J-J-D	0.2130	0.2951	0.3804	0.4838
J-J-O	0.2663	0.3477	0.3915	0.4963

Table 6 MAP of overall K-K runs

(a) Average and median

	Average		Median	
	Rigid	Relax	Rigid	Relax
K-K	0.4109	0.4402	0.4431	0.4699
K-K-T	0.4271	0.4582	0.4588	0.4934
K-K-D	0.3869	0.4149	0.3727	0.3992
K-K-O	0.4171	0.4457	0.4694	0.5004

(b) Min and max

	Min		Max	
	Rigid	Relax	Rigid	Relax
K-K	0.0000	0.0000	0.5825	0.6212
K-K-T	0.2821	0.3136	0.5078	0.5361
K-K-D	0.2297	0.2587	0.4685	0.5097
K-K-O	0.0000	0.0000	0.5825	0.6212

Table 7 MAP of overall E-E runs

(a) Average and median

	Average		Median	
	Rigid	Relax	Rigid	Relax
E-E	0.3102	0.3908	0.3161	0.4042
E-E-T	0.2963	0.3767	0.3145	0.3954
E-E-D	0.2895	0.3676	0.3026	0.3859
E-E-O	0.3518	0.4357	0.3573	0.4423

(b) Min and max

	Min		Max	
	Rigid	Relax	Rigid	Relax
E-E	0.0342	0.0483	0.4000	0.4962
E-E-T	0.0802	0.1032	0.3576	0.4512
E-E-D	0.0342	0.0483	0.3469	0.4368
E-E-O	0.2864	0.3627	0.4000	0.4962

7.1.5 Remarks. The average values of MAP of all C-C runs, all J-J runs, all K-K runs and all E-E runs based on rigid relevance are 0.1985, 0.3258, 0.4109, and 0.3102, respectively (see Table 8, 10, 12 and 14). It seems that Chinese monolingual runs are more difficult than the other languages and Korean monolingual runs easier.

7.2 BLIR runs on Chinese document set

7.2.1 J-C runs. In total, 8 J-C runs were submitted by only 2 groups. Table 8 shows the best runs of each group (only D-runs based on rigid relevance). While IFLAB uses standard bilingual resources (dictionary and corpus-based), OKI apply a pivot language method. As shown in Table 16, the MAP values of J-C runs are not high.

Table 8 Best runs (J-C, Rigid, D-runs)

Run-ID	MAP
IFLAB-J-C-D-01	0.0548
OKI-J-C-D-04 (pivot)	0.0404

7.2.2 K-C runs. Only one group submitted search results of K-C runs (5 runs were submitted). The best run is KLE-K-C-D-01, of which MAP value is 0.1447.

7.2.3 E-C runs. In total, 12 E-C runs were submitted by 3 groups. Table 9 shows the best runs of each group (only D-runs based on rigid relevance). As similar with J-C runs, the performance is low.

Table 9 Best runs (E-C, Rigid, D-runs)

Run-ID	MAP
TJUCN-E-C-D-01	0.0663
OKI-E-C-D-04	0.0481
ISCAS-E-C-D-03	0.0017

7.2.4 Remarks. The values of MAP in J-C and E-C runs are low, and there seems to be a room for further research efforts. Meanwhile, KLE runs show better performance in K-C retrieval. Values of MAP of the best J-C runs, the best K-C run and the best E-C run (D-runs) are 0.0548, 0.1447, and 0.0663, respectively, and these MAP values are 16.8%, 44.4%, 20.4% of that of the best C-C run (0.3255, rigid and D-runs), respectively.

7.3 BLIR runs on Japanese document set

7.3.1 C-J runs. In total, 18 C-J runs were submitted by 5 groups. Table 10 shows the best runs of each

group (only D-runs based on rigid relevance). While TSB and NII use MT systems, the other three groups, BRKLY, OKI and tlrrd, adopt pivot language approach. Only performance of BRKLY using MT system-based pivot approach is comparable with non-pivot approach (especially, the BRKLY run outperforms the NII run).

7.3.2 K-J runs. In total, 13 K-J runs were submitted by 4 groups. Table 11 shows the best runs of each group (only <DESC>-only run based on rigid relevance). While KLE and NII use direct bilingual resources, the other two groups, BRKLY and tlrrd, adopt pivot language approach. Performance of pivot language approach does not reach to the level of that of non-pivot approach.

7.3.3 E-J runs. In total, 15 E-J runs were submitted by 4 groups. Table 12 shows the best runs of each group (only D-runs based on rigid relevance). While TSB and NII use direct MT systems, the other two groups, OKI and BRKLY, adopt pivot language approach. The OKI's run outperforms runs by non-pivot, direct translation approach.

7.3.4 Remarks. Values of MAP of the best C-J runs, the best K-J run and the best E-J run (D-runs) are 0.2309, 0.2935, and 0.2674, respectively, and these MAP values are 60.7%, 77.2%, 70.3% of that of the best J-J run (0.3804, rigid and D-runs), respectively. Pivot language approach shows comparable performance with non-pivot, direct translation approach in C-J and E-J runs.

Table 10 Best runs (C-J, Rigid, D-runs)

Run-ID	MAP
TSB-C-J-D-03	0.2309
BRKLY-C-J-D-03	0.1904
NII-C-J-D-02	0.1455
OKI-C-J-D-04	0.1088
tlrrd-C-J-D-02	0.0544

Table 11 Best runs (K-J, Rigid, D-runs)

Run-ID	MAP
KLE-K-J-D-01	0.2935
NII-K-J-D-02	0.1894
BRKLY-K-J-D-03	0.1402
tlrrd-K-J-D-02	0.0964

Table 12 Best runs (E-J, Rigid, D-runs)

Run-ID	MAP
OKI-E-J-D-04	0.2674
TSB-E-J-D-01	0.2672
NII-E-J-D-02	0.2533
BRKLY-E-J-D-02	0.1874

7.4 BLIR runs on Korean document set

7.4.1 C-K runs. In total, 8 C-K runs were submitted by only 2 groups. Table 13 shows best runs of each group (only D-runs based on rigid relevance). The pircs run adopts pivot language approach, and its performance is lower than the KLE run using direct translation approach.

7.4.2 J-K runs. In total, 8 J-K runs were submitted by only 2 groups. Table 14 shows the best runs of each group (only D-runs based on rigid relevance).

7.4.3 E-K runs. In total, 7 E-K runs were submitted by only 2 groups. Table 15 shows the best runs of

each group (only D-runs based on rigid relevance).

7.4.4 Remarks. Values of MAP of the best C-K runs, the best J-K run and the best E-K run (D-runs) are 0.3973, 0.3984, and 0.3249, respectively, and these MAP values are 84.8%, 85.0%, 69.3% of that of the best K-K run (0.4685, rigid and D-runs), respectively.

Table 13 Best runs (C-K, Rigid, D-runs)

Run-ID	MAP
KLE-C-K-D-01	0.3973
pircs-C-K-D-02	0.2471

Table 14 Best runs (J-K, Rigid, D-runs)

Run-ID	MAP
KLE-J-K-D-01	0.3984
IFLAB-J-K-D-01	0.2363

Table 15 Best runs (E-K, Rigid, D-runs)

Run-ID	MAP
pircs-E-K-D-02	0.3249
KLE-E-K-D-01	0.0981

7.5 BLIR runs on English document set

7.5.1 C-E runs. In total, 24 C-E runs were submitted by 7 groups. Table 16 shows the best runs of each group (only D-runs based on rigid relevance). The TSB run was executed using Japanese as a pivot (by two MT systems). The TBS and pircs employ MT system and the other 5 runs are based on bilingual dictionaries.

7.5.2 J-E runs. In total, 23 J-E runs were submitted by 6 groups. Table 17 shows the best runs of each group (only D-runs based on rigid relevance). The TBS run and the OKI run were executed using MT system (OKI also employs parallel corpus and a bilingual dictionary) and for the other 4 runs, no MT system seems to be used.

7.5.3 K-E runs. In total, 8 K-E runs were submitted by only 2 groups. Table 18 shows the best runs of each group. (only D-runs based on rigid relevance)

7.5.4 Remarks. Values of MAP of the best C-E runs, the best J-E run and the best K-E run (D-runs) are 0.2238, 0.3340, and 0.2250, respectively, and these MAP values are 64.5%, 96.2%, 64.9% of that of the best E-E run (0.3469, rigid and D-runs), respectively. Especially, the top E-J run shows almost same performance with the best E-E monolingual run.

Table 16 Best runs (C-E, Rigid, D-runs)

Run-ID	MAP
pircs-C-E-D-02	0.2238
TSB-C-E-D-01	0.2183
RMIT-C-E-D-04	0.1918
UCNTC-C-E-D-02	0.1758
JSCCC-C-E-D-01	0.1575
OKI-C-E-D-04	0.1265
AILAB-C-E-D-01	0.0412

Table 17 Best runs (J-E, Rigid, D-runs)

Run-ID	MAP
TSB-J-E-D-01	0.3340
OKI-J-E-D-04	0.2813
JSCCC-J-E-D-02	0.2620
IFLAB-J-E-D-01	0.2225
FORES-J-E-D-01	0.0775
UENIS-J-E-D-02	0.0075

Table 18 Best runs (J-E, Rigid, D-runs)

Run-ID	MAP
KUNLP-K-E-D-02	0.2250
KLE-K-E-D-01	0.1876

7.6 MLIR

In the case of MLIR on CJE collection, 9 C-CJE runs submitted by 2 groups, 5 J-CJE runs by one group, and 15 E-CJE runs by 3 groups. For MLIR on larger CJKE collection, 3 J-CJKE runs submitted by one group and 5 E-CJKE runs by one group. Table 19 shows the best runs of each group by run type (only D-runs based on rigid relevance).

Table 19 Best runs (MLIR, Rigid, D-runs)

Run-ID	MAP
C-CEJ: OKI-C-CEJ-D-04	0.0923
NTU-C-CJE-D-01	0.0521
J-CEJ: OKI-J-CEJ-D-04	0.1566
E-CJE: UniNE-E-CJE-D-02	0.1604
OKI-E-CEJ-D-04	0.1588
UCNTC-E-CJE-D-02	0.0877
J-CJKE: IFLAB-J-CJKE-D-01	0.1296
E-CJKE: UniNE-E-CJKE-D-03*	0.1766

*It should be noted that UniNE-E-CJKE-D-03 includes a search on the Korean collection based of the DNC topic sections.

8. Concluding remarks

It seems that more various approaches or techniques are investigated at this task than before, e.g., a variety of Chinese indexing methods, transliteration techniques, Web-based solutions for out-of-vocabulary problem, applications of LM, new challenges to query expansion, and so on. However, there is a room for further research efforts for improving Chinese information retrieval in comparison with Korean and Japanese IR. It turns out at this task that performance of Korean IR is relatively high. This may be partly because Korean text includes white spaces as delimiters. We need further investigation on CLIR between CJK languages.

References

- [1] Kando, Noriko. Overview of the Second NTCIR Workshop. In *Proceedings of the Second NTCIR Workshop on Evaluation of Chinese & Japanese Text Retrieval and Text Summarization*, pages 51-72, Tokyo, 2001.
- [2] Chen, Kuang-hua & Hsin-His Chen. The Chinese Text Retrieval tasks of NTCIR Workshop 2. In *Proceedings of the Second NTCIR Workshop on Evaluation of Chinese & Japanese Text Retrieval and Text Summarization*, pages 51-72, Tokyo, 2001.
- [3] Chen, Kuang-hua, Hsin-Hsi Chen, N. Kando, K. Kuriyama, S. Lee, S. H. Myaeng, K. Kishida, K. Eguchi & H. Kim. Overview of CLIR task at the third NTCIR Workshop. In *Working Notes of the Third NTCIR Workshop Meeting, Cross-Lingual Information Retrieval Task*, pages 1-38, 2002.