Overview of the Fifth NTCIR Workshop

Noriko Kando

National Institute of Informatics

http://research.nii.ac.jp/ntcir/

kando (at) nii. ac. jp

NTCIR Workshop is :

A series of evaluation workshops designed to enhance research in **information access** technologies by providing infrastructure of large-scale evaluation.

Project started late 1997, Once per 1½ years

1st : Nov.1,1998- Sept.1,1999

2nd : June,2000– March,2001

3rd : Sept 2001- Oct 2002

4th: Apr 2003 – June 2004

5th: Oct 2004 – Dec 2005

* **N**ii **T**est **C**ollection for **I**nformation **R**etrieval systems

* Co-sponsored by NII and MEXT Grant-in-Aid on Informatics

Focus of NTCIR

Lab-type IR Test

Asian Languages/cross-language

Variety of Genre

Parallel/comparable Corpus

New Challenges

Intersection of IR + NLP

To make information in the documents more usable for users!

Realistic eval/user task

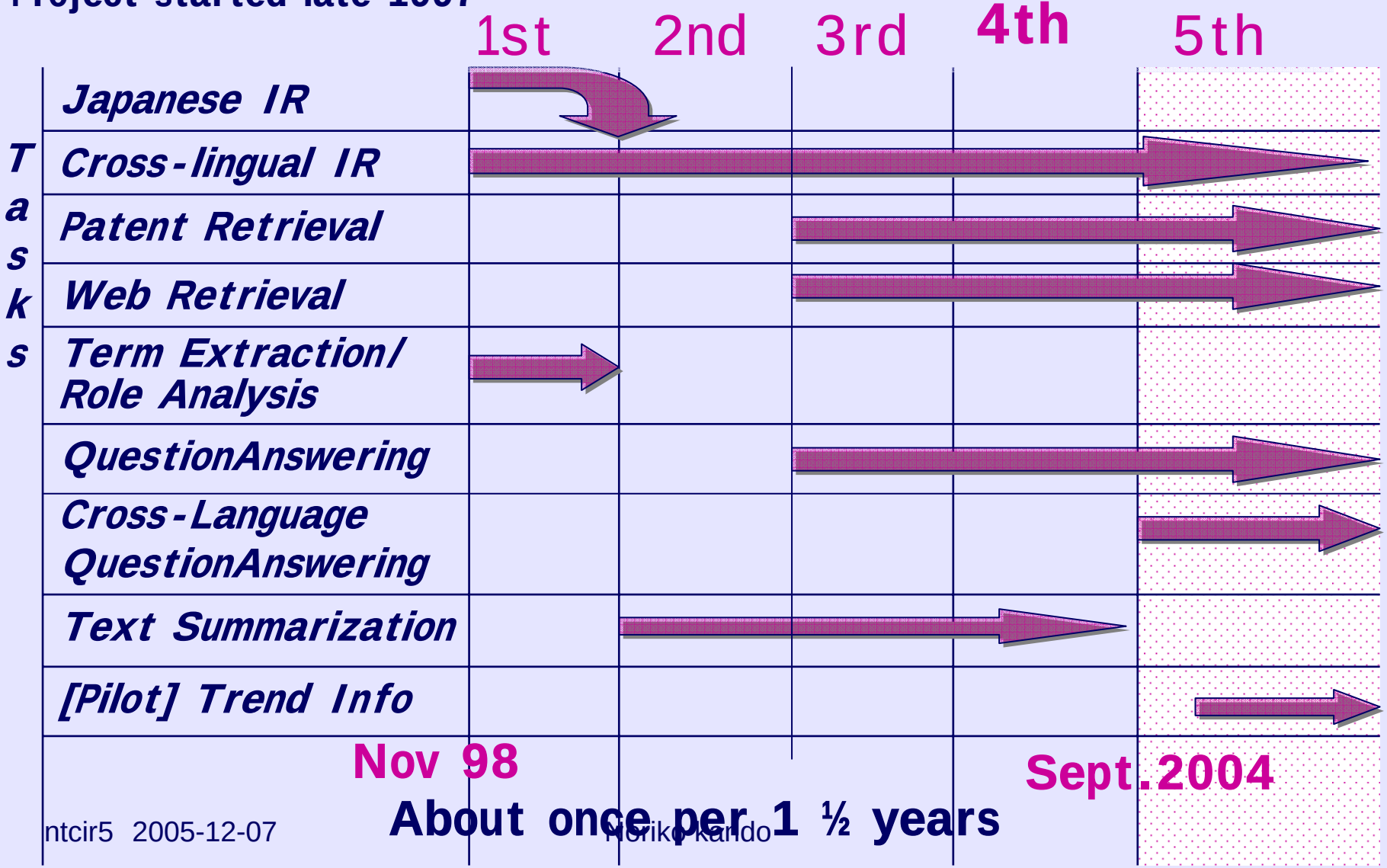
Forum for Researchers

Idea Exchange

Discussion/ Investigation on Evaluation methods/metrics

Tasks (Research Areas) of NTCIR Workshops

Project started late 1997



NTCIR-5 Tasks

- **CLIR:** focus: NE, OOV, news docs 2000-2001CJK
- **CLQA:** (Pilot, New)
  - On C doc: E-C, C-C / - On JE doc: E-J, J-E, C-E
- **Patent Retrieval: 10 yr patent fulltext ca45GB**
  - Document Retrieval (Invalidity Search) 1200+ topics
  - Passage Retrieval
  - Text Categorization to F-terms (good granularity for patent map)
- **QAC:** Series of Questions (J-J)
- **WEB:**
  - Navigational Retrieval, New **1.36TB** docs
  - Query Term Expansion
- [Pilot] **Must:** MUltimodal Summarization for Trend information, extract numeric information from a set of documents, and visualize them to show their trends

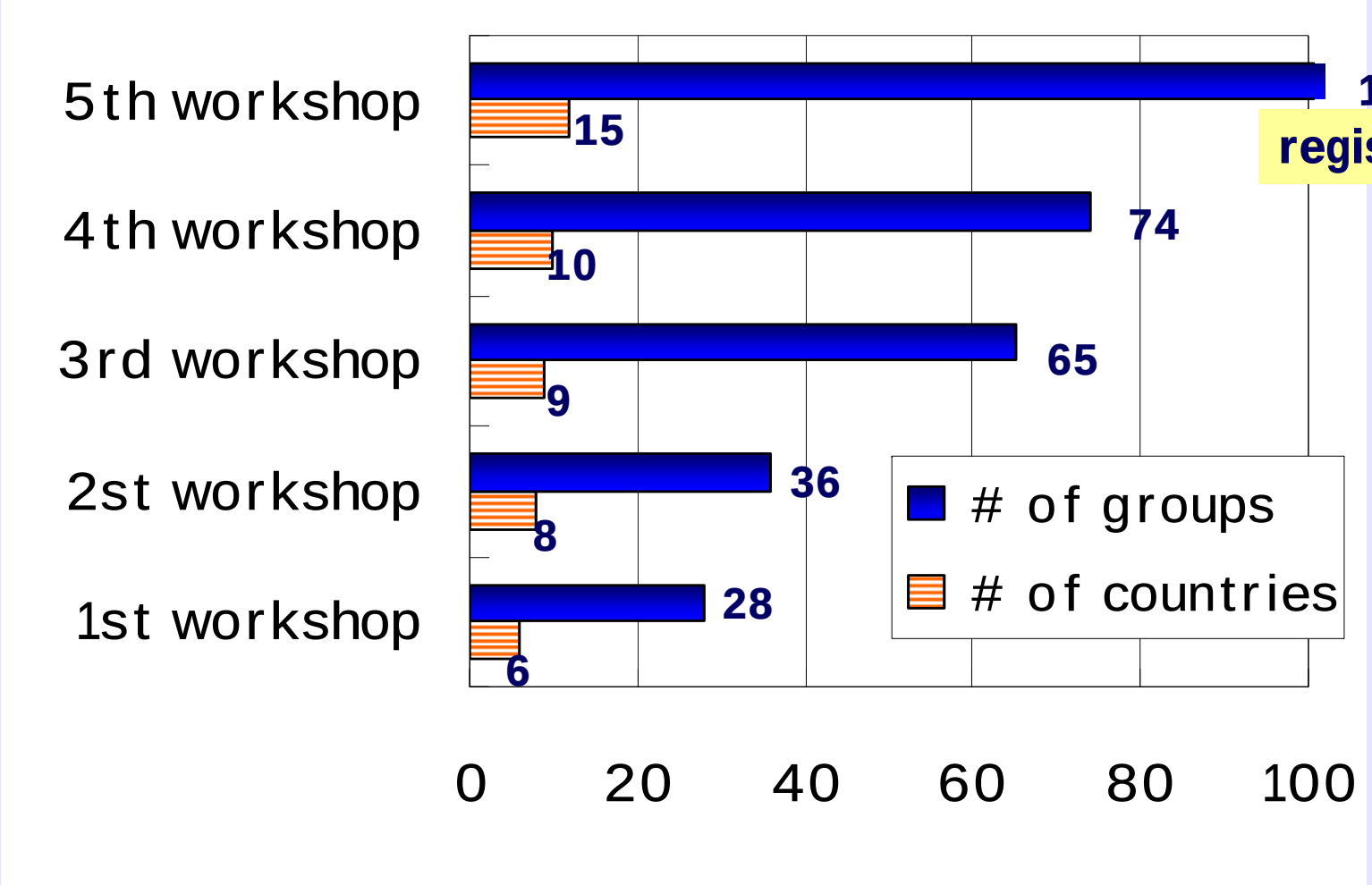
Schedule for NTCIR-5

[TASK]
Sept-Dec 2004: Document Release
April-July, 2005:Formal Run
1 Sept 2005: Evaluation Results Return (excpt CLIR)
15 Oct 2005: Paper Submission
6-9 Dec 2005: Meeting, at NII
*Proceedings will be published at the Conference.

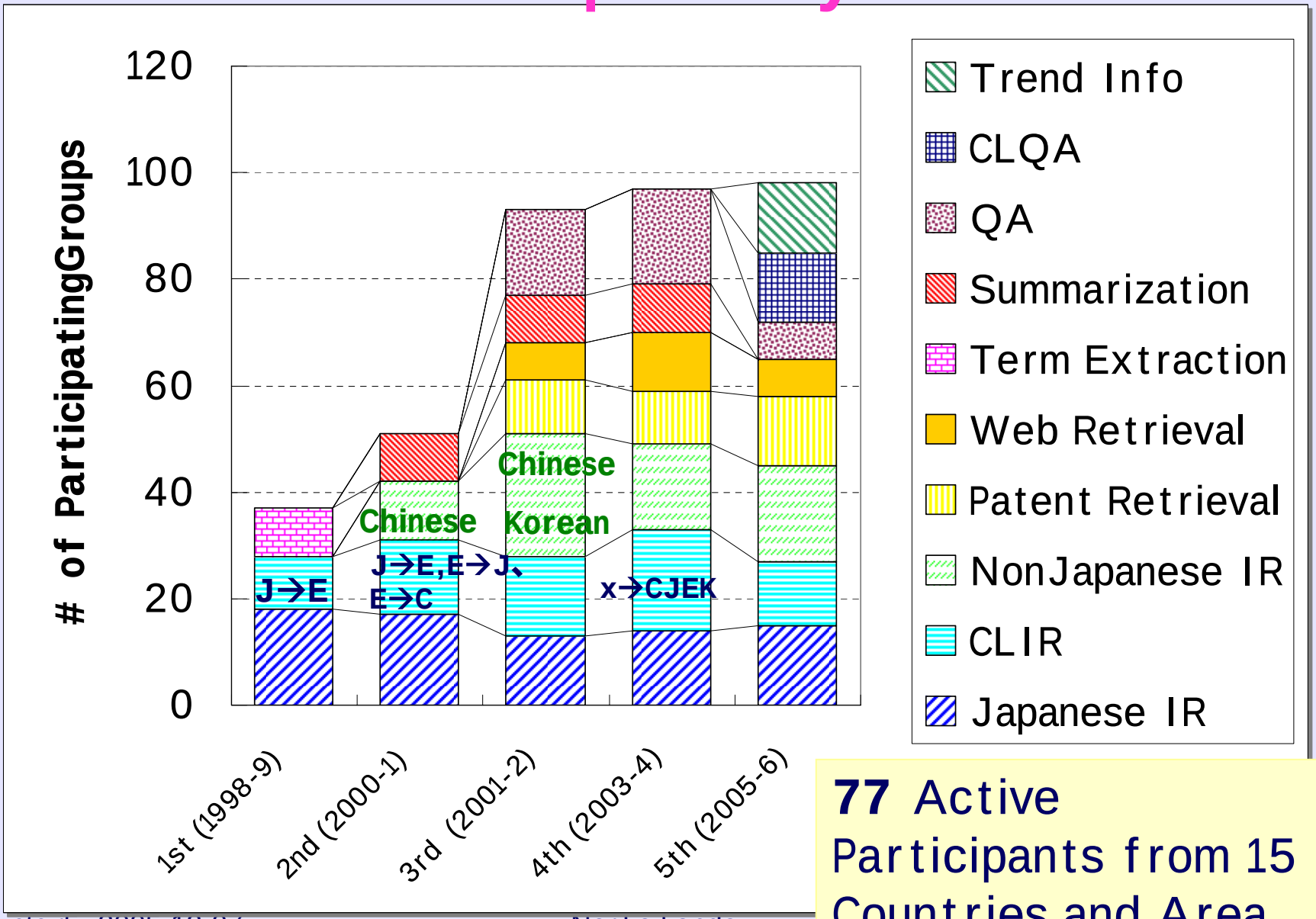
[**Open Submission**]
1 Oct 2005: Paper Due
1 Nov 2005: Late Breaking Short paper Due
15 Nov 2005: Notification

[Pilot WS: **MuST**] Workshop in March, 2006, Tokyo

NTCIR workshop: Number of Participating Groups



Number of Participants by Tasks



77 Active Participants from 15 Countries and Area

Active Participants

[CLIR]
- DAEDALUS Data/ Decisions and Language
- Fu Jen Catholic Univ
- Hummingbird
- Institute for Infocomm Research
- Institute of Software; Chinese Academy of Sciences
- National Cheng Kung Univ
- National Institute of Information and Communications Technology
- NII / Univ of Hildesheim
- Oki Electric Industry
- Pohang Univ of Science and Engineering
- Queens College; CUNY
- Queensland Univ of Technology
- RMIT Computer Science & IT
- Ryukoku Univ / National Institute of Information and Communications Technology
- The Hong Kong Polytechnic Univ
- TLR R & D
- Toshiba Corporate R&D Center
- Tsinghua Univ/ Central China Normal Univ
- Univ of Amsterdam
- Univ of California, Berkeley
- Univ of Helsinki
- Univ of Neuchatel
- Unversity of NorthTexas
- Yahoo

[CLQA]
- Academia Sinica; Taiwan
- ATR Spoken Language Translation Research Laboratoires
- Carnegie Mellon Univ
- National Cheng Kung Univ
- Institute of Computing Technology; Chinese Academy of Sciences/ Dublin City Univ
- National Institute of Information and Communications Technology
- National Taiwan Ocean Univ
- NTT CS Lab.
- Queens College; CUNY
- Toyohashi Univ of Technology/ Univ of Tsukuba/ Nagoya Univ
- Univ of Limerick
- Univ of North Texas
- Yokohama National Univ

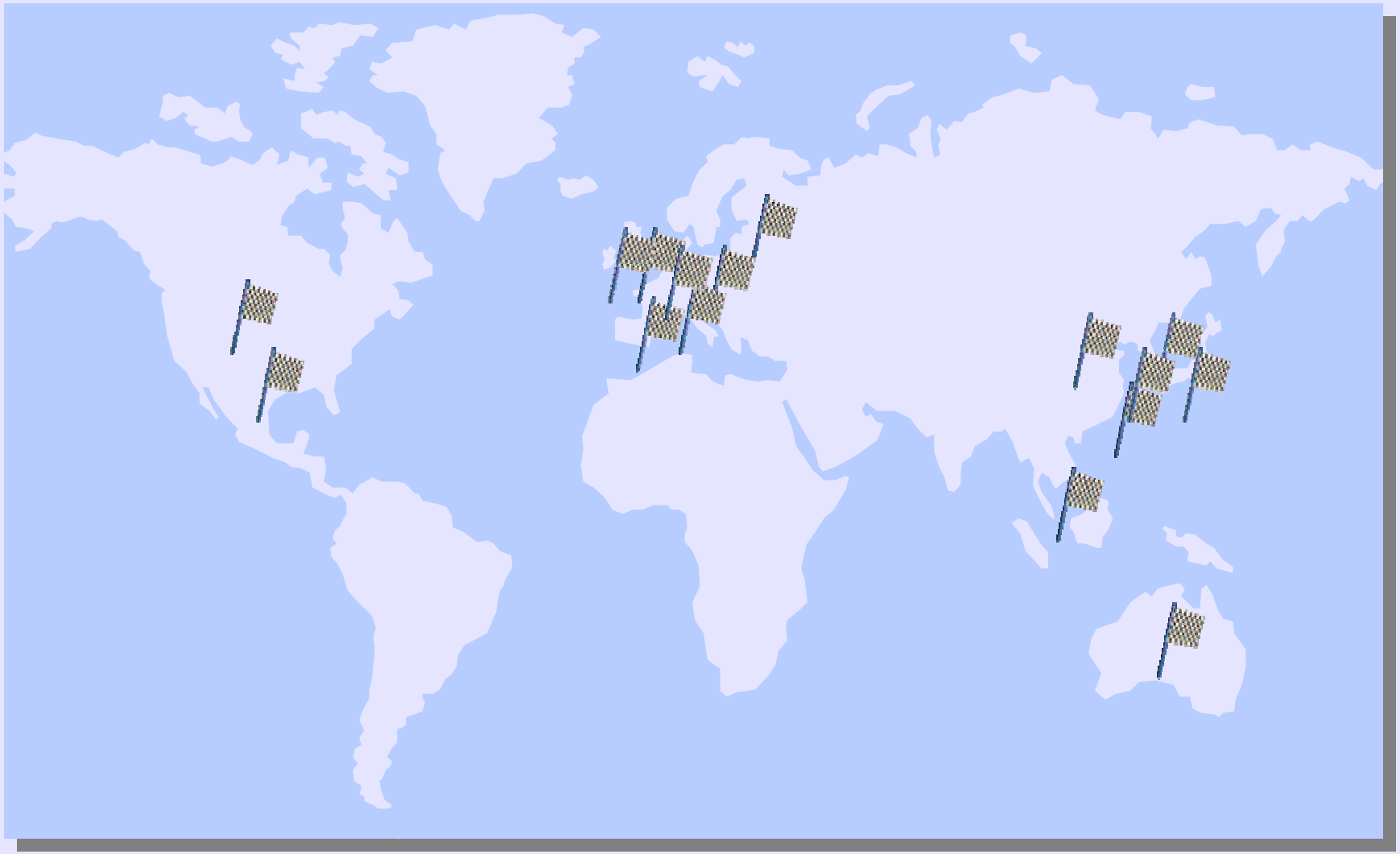
[PATENT]
- IBM Research
- Fuji Xerox
- Fujitsu Laboratories
- Hitachi
- Justsystem Corporation
- KAIST
- National Institute of Information and Communications Technology
- NTT DATA CORPORATION
- Pohang Univ of Science and Technology
- RICOH
- Toyohashi Univ of Technology
- Univ of Tsukuba
- WebGenie Information Ltd.

[QAC]
- Hokkaido Univ
- Iwate Prefectural Univ
- Nagaoka Univ of Technology
- National Institute of Information and Communications Technology
- Ritsumeikan Univ
- Toyohashi Univ of Technology/ Univ of Tsukuba/ Nagoya Univ
- Yokohama National Univ

[WEB]
- NEC Corporation
- Justsystem Corporation
- NII; Univ. of Tokyo/ KYA group
- Osaka Kyoiku Univ
- RICOH
- Univ of Aizu
- Univ of Tsukuba

[MuST]
- Dai Nippon Printing Co.,Ltd,
- Hiroshima City Univ
- Justsystem Corporation
- Keio Univ
- National Institute of Information and Communications Technology
- NTT CS Lab.
- Ochanomizu Univ
- OKI
- Osaka Prefecture Univ
- The Graduate School of Nataural Science and Technology
- Tokyo Metoropolitan Univ Faculty of System Design
- Toyohashi Univ of Technology
- Yokohama National Univ

Geographical Distribution of Participants



Geographical Distribution of Participants

NTCIR Workshop 5 (2004-2005) Organizers

Program chair: Noriko Kando, NII

+CLIR
Hsin-Hsi Chen, NTU
Kuang-hua Chen, NTU
Kazuaki Kishida, Surugadai U
Kazuko Kuriyama, Shirayuri U
Sukhoon Lee, NCU
Sung Hyon Myaeng, IIU
Noriko Kando, NII

+CLQA
Kuang-hua Chen, NTU
Chuan-Jie Lin , Nat Taiwan Ocean U
Yutaka Sakaki, ATR

+PATENT
Atsushi Fujii, Tsukuba U
Makoto Iwayama, Hitachi/TITEC
Noriko Kando, NII

+QA
Junichi Fukumoto, Ritsumeikan U
Tsuneaki Kato, U Tokyo
Fumito Masui, Mie U

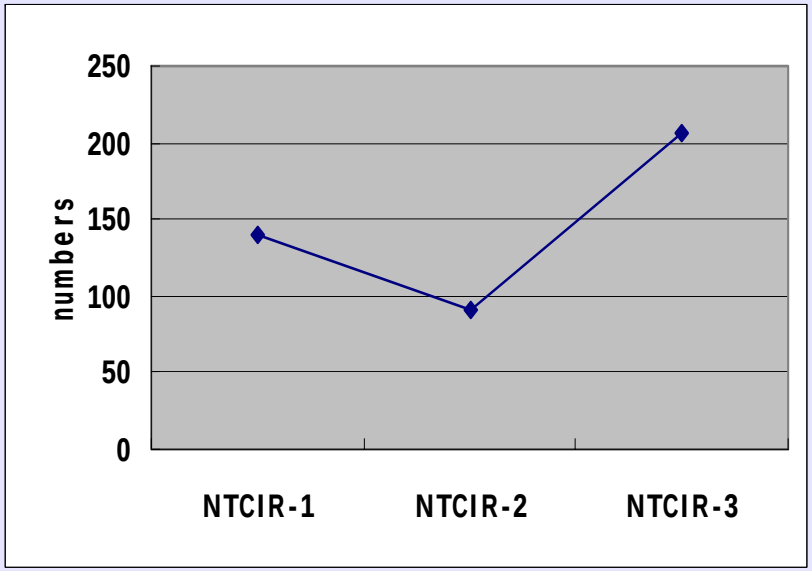
+WEB
Keizo Oyama, NII
Masao Takaku, NII

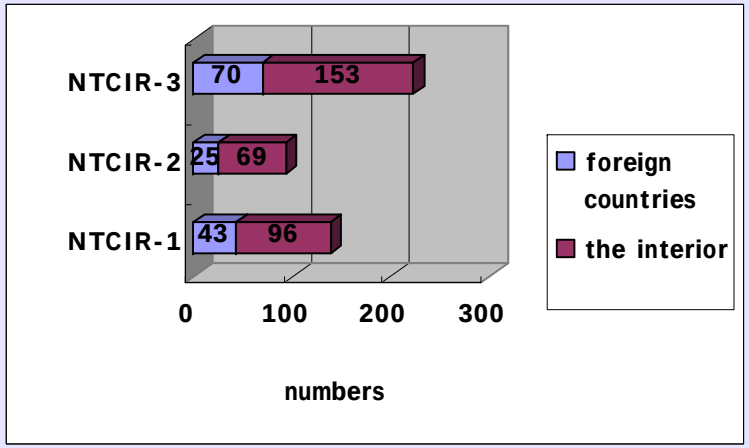
NTCIR test collections

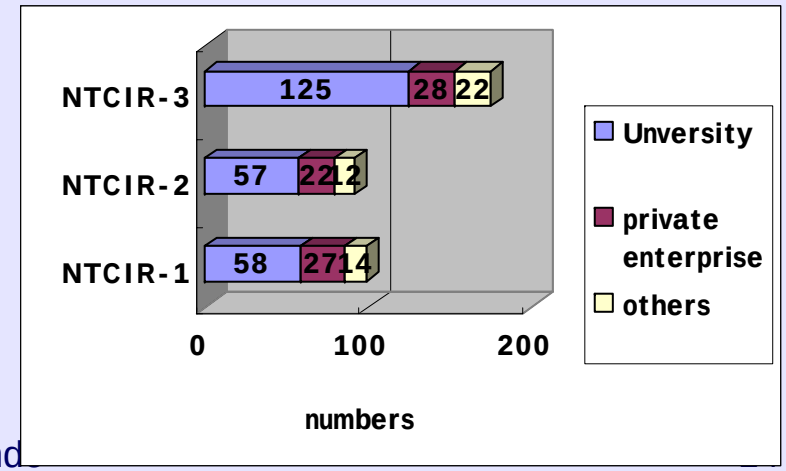
Collection	task	Documents			topic./Q	Relevance/
		Genre	Size	Language	Language	Answer
NTCIR-1	IR	Academic	577MB	JE	J	3
CIRB010	IR	News	132MB	Ct	CtE	4
NTCIR-2	IR	Academic	800MB	JE	JE	4
NTCIR-2 Summ	Summ	News	180 docs	J	J	
NTCIR-3 CLIR	IR	News	884MB	CtKJE	CtKJE	4
NTCIR-3 PATENT	IR	Patent	18GB(+5GB)	J(JE)	CsCtKJE	3
NTCIR-3 QA	QA	News	282MB	J	J(E)	exact
NTCIR-3 Summ	Summ	News	60 docs+50 sets	J	—	
NTCIR-3 WEB	IR	WEB	100GB	Multiple	J(E)	4+relative
NTCIR-4 CLIR	IR	News	ca 3GB	CtKJE	CtKJE	4
NTCIR-4 PATENT	IR	Patent	45GB	J(JE)	CsCtKJE	3
NTCIR-4 QA	QA	News	776MB	J	J(E)	4
NTCIR-4 Summ	Summ	News	30 sets	J	—	
NTCIR-4 WEB	IR	WEB	100GB	Multiple	J(E)	

Ct：Traditional Chinese、Cs：Simplified Chinese、K：Korean、J：Japanese、E：English

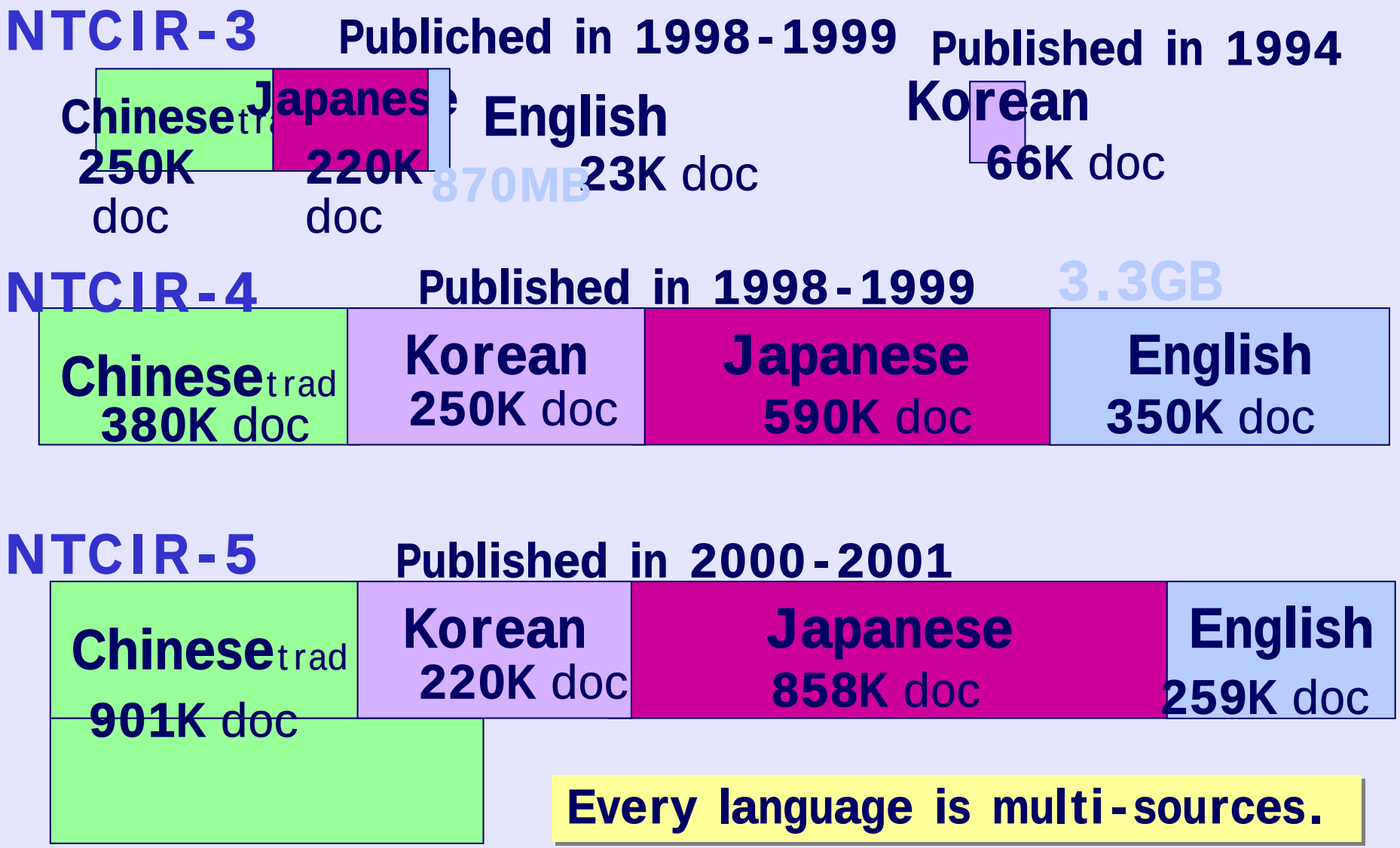
Situation on the Data Distribution of Research Purpose Use of NTCIR Test Collections



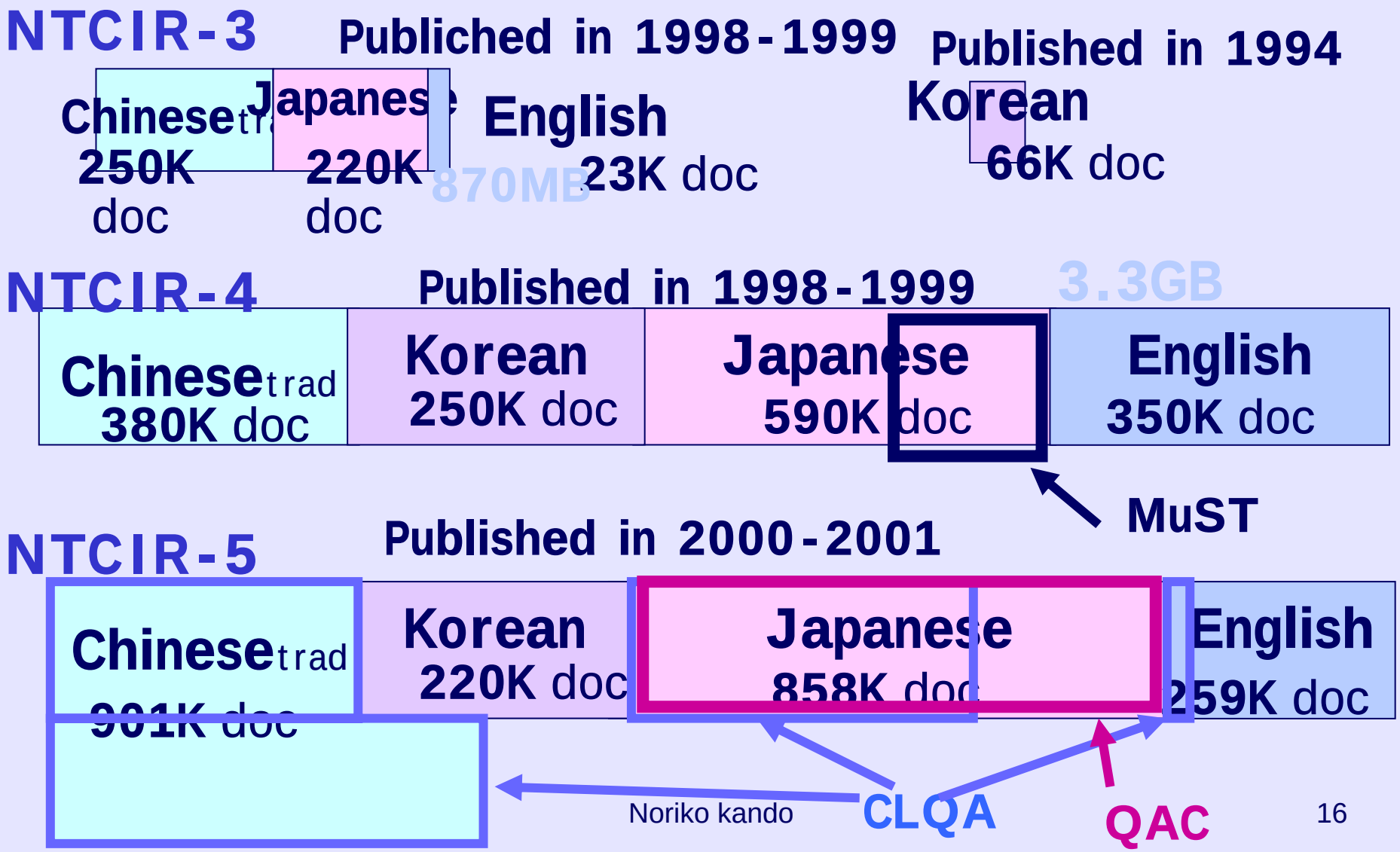




News Documents NTCIR-5

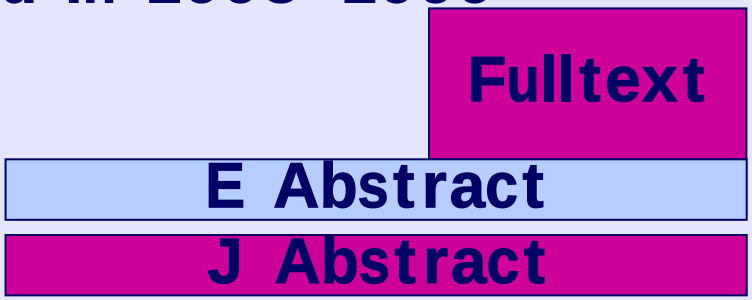


News Documents NTCIR-5



Patents Documents NTCIR-5

NTCIR-3 Filed in 1998-1999



31 Q

NTCIR-4 Filed in 1993-1997

34+69 Q

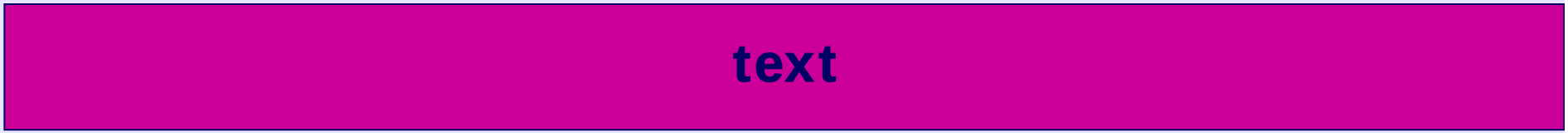
NTCIR-5 Filed in 1993-2002 34+1189 Q

WEB Documents NTCIR-5

NTCIR-3 & -4 NW100GB-01 47 Q

100GB

NTCIR-5 NW1000GB-04 269+847 Q



1.36TB

What were New to NTCIR-4

- Open Submission Session
- ACM-TALIP Special Issue Recommendation
- Open Attendance
- Submission Raw Data
- Online Working Notes and Slides

What's New to NTCIR-5

- Open Submission >>>> continued
- ACM-TALIP Special Issue (strategy change necessary), but **Special Issue on Patent at IP&M**
- Open Attendance >>>>continued
- Submission Raw Data >>>>continued
- Online Working Notes and Slides >>>>

  Proceedings at Conference Only (No working notes)
  Online:http://research.nii.ac.jp/ntcir/workshop/OnlineProceedings5/index.html
- **Pilot tasks** and **feasibility studies** MuST, "why" question

Acknowledgment

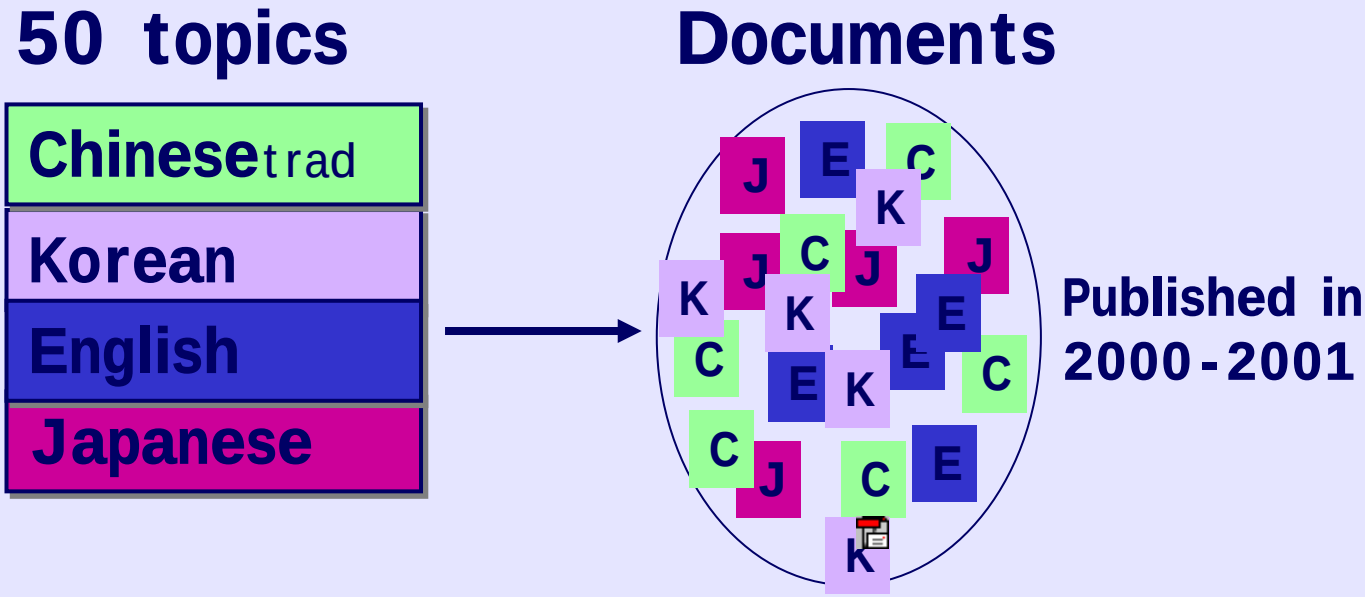
- Central Daily News
- China Daily News
- China Times Inc.
- ChosunIlbo
- Hankooki.com
- Industrial Property Cooperation Center
- Japan Parent Office
- Japan Patent Information Organization

Korea Economic Daily
Linguistic Data Consortium
Mainichi Newspaper
Nippon Database Kaihatsu, Co. Ltd.
NTT
NRI Cyber Patent
PATOLIS
the Sing Tao Group
Taiwan News
Tokyo Univ
UDN.COM
Wisers Information Ltd.
Yomiuri Shinbun

Cross-Language Information Retrieval (CLIR) Task

Task Organizers
Kazuaki Kishida*, Kuang-hua Chen, Sukhoon Lee,
Hsin-Hsi Chen, Koji Eguchi, Noriko Kando
Kazuko Kuriyama, Sung Hyon Myaeng

NTCIR-5 CLIR



- **Short Q**: D-only and T-only are mandatory
- **Background info** of search requests
- Balance btw **topic-types**:

- Forcus: NE, OOV
- proper nouns vs without PN
- domestic/regional/international

Design of CLIR Task

- Subtasks
  - Multilingual CLIR (MLIR) : e.g., C - CJKE
  - Bilingual CLIR (BLIR): e.g., C - J
  - Single Language IR (SLIR): e.g., C - C

  Languages
  - Chinese (C), Japanese (J), Korean (K), English (E)
- Relevance Judgments – 4 grades
  - Highly Relevant (S), Relevant (A), Partial Relevant (B), Non-Relevant (C)

Result Submission

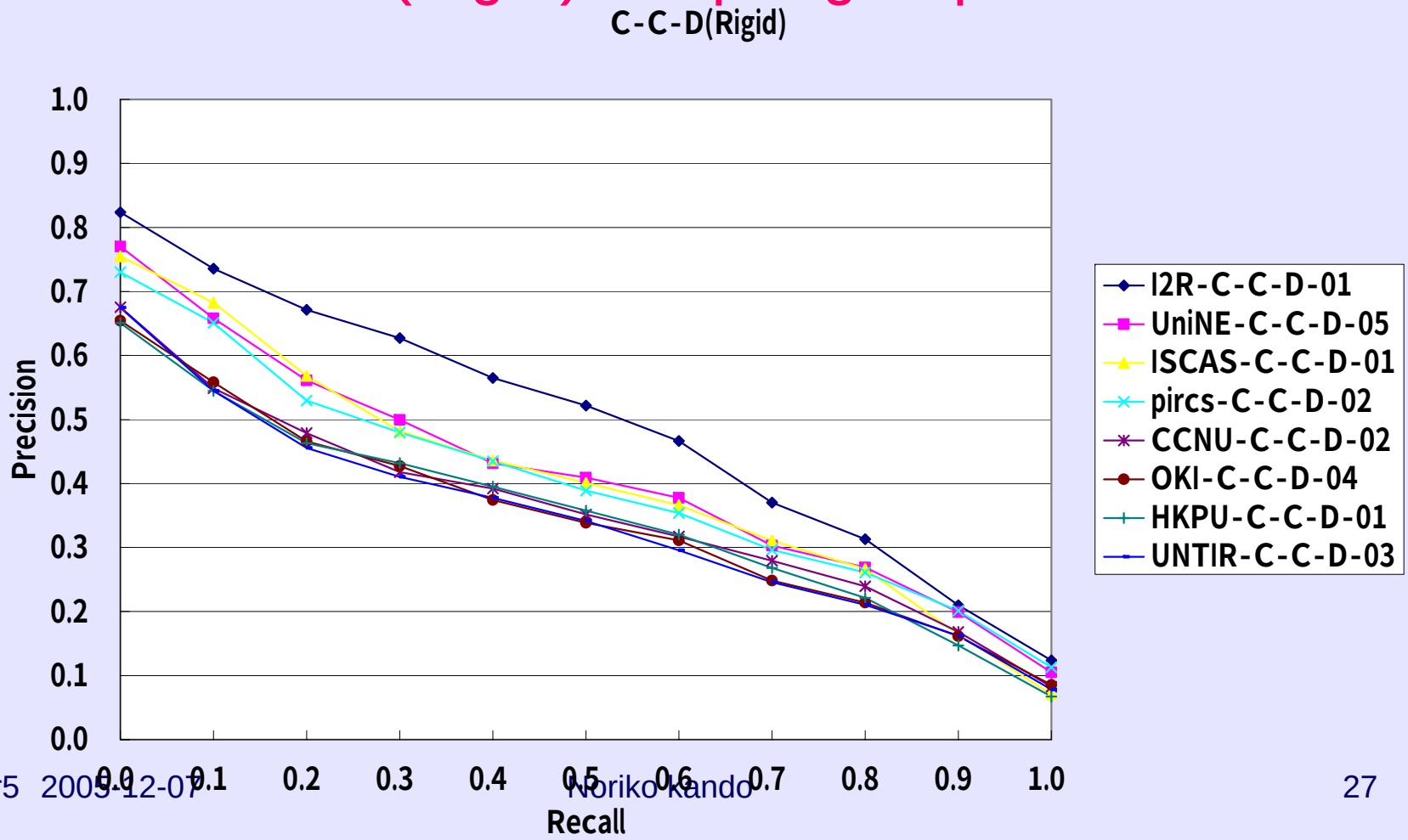
- 24 groups submitted results
  - From Australia, Canada, China PRC, Finland, Germany, Hong Kong, Japan, Korea, Netherlands, Singapore, Spain, Switzerland, Taiwan, USA (14 countries and areas)

Techniques Used (NTCIR-5)

- IR Models: WIN, Logistic Reg, pircs, others vsm, okapi, LM, etc.
- Indexing: bigram vs word vs others, hybrid
- Mostly “Query Trans”, but one “Doc Trans”
- Translation disambiguation w/ WEB w/target doc
- Out-of-vocabulary (OOV) problem
  - NE identification
  - Use of Web
  - Transliteration - Cognate
- Query expansion techniques
  - Selective application PRF, Bounce & Throw
  - Clustering

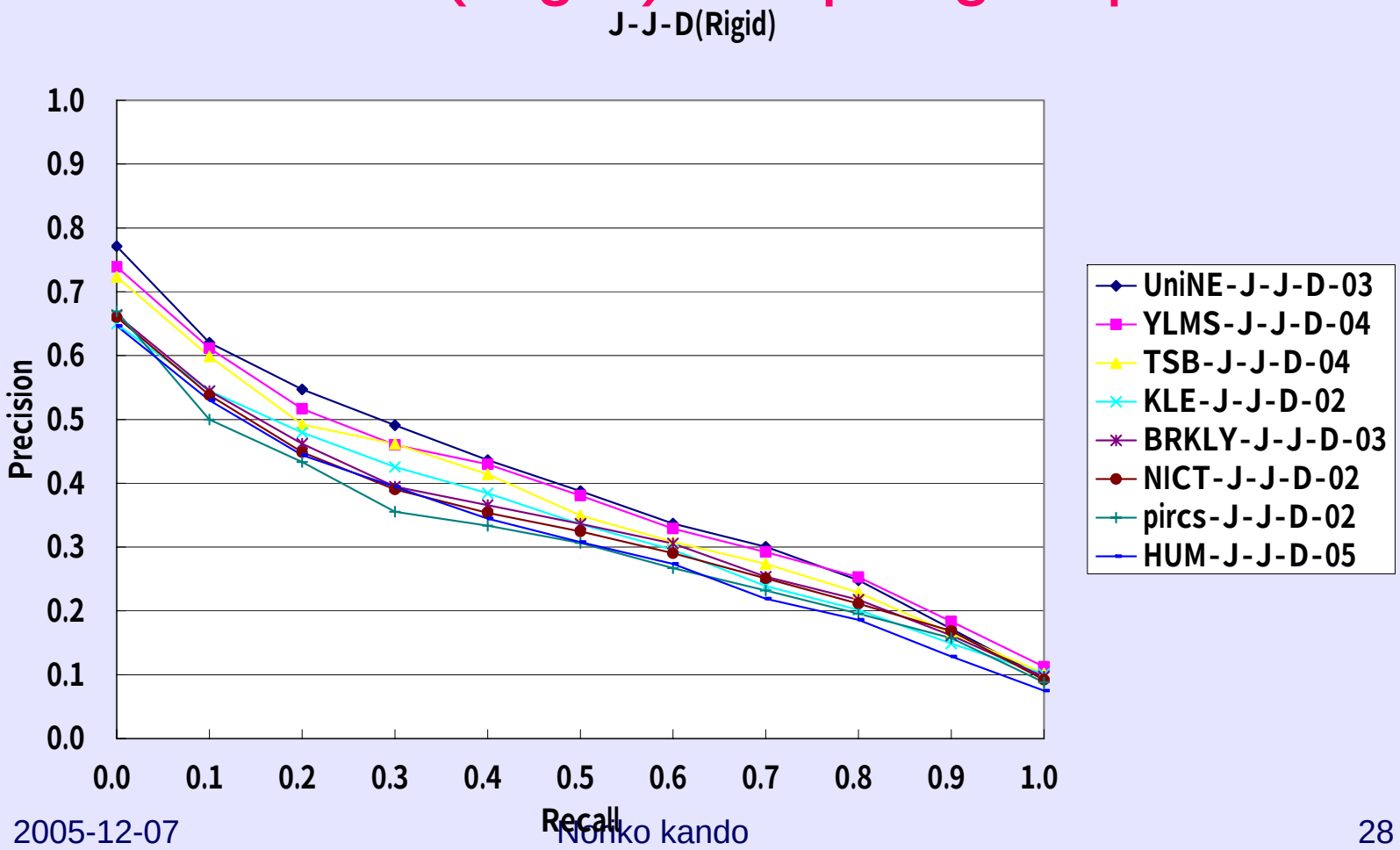
Evaluation (2)

- SLIR: C-C-D (Rigid) – top 8 groups



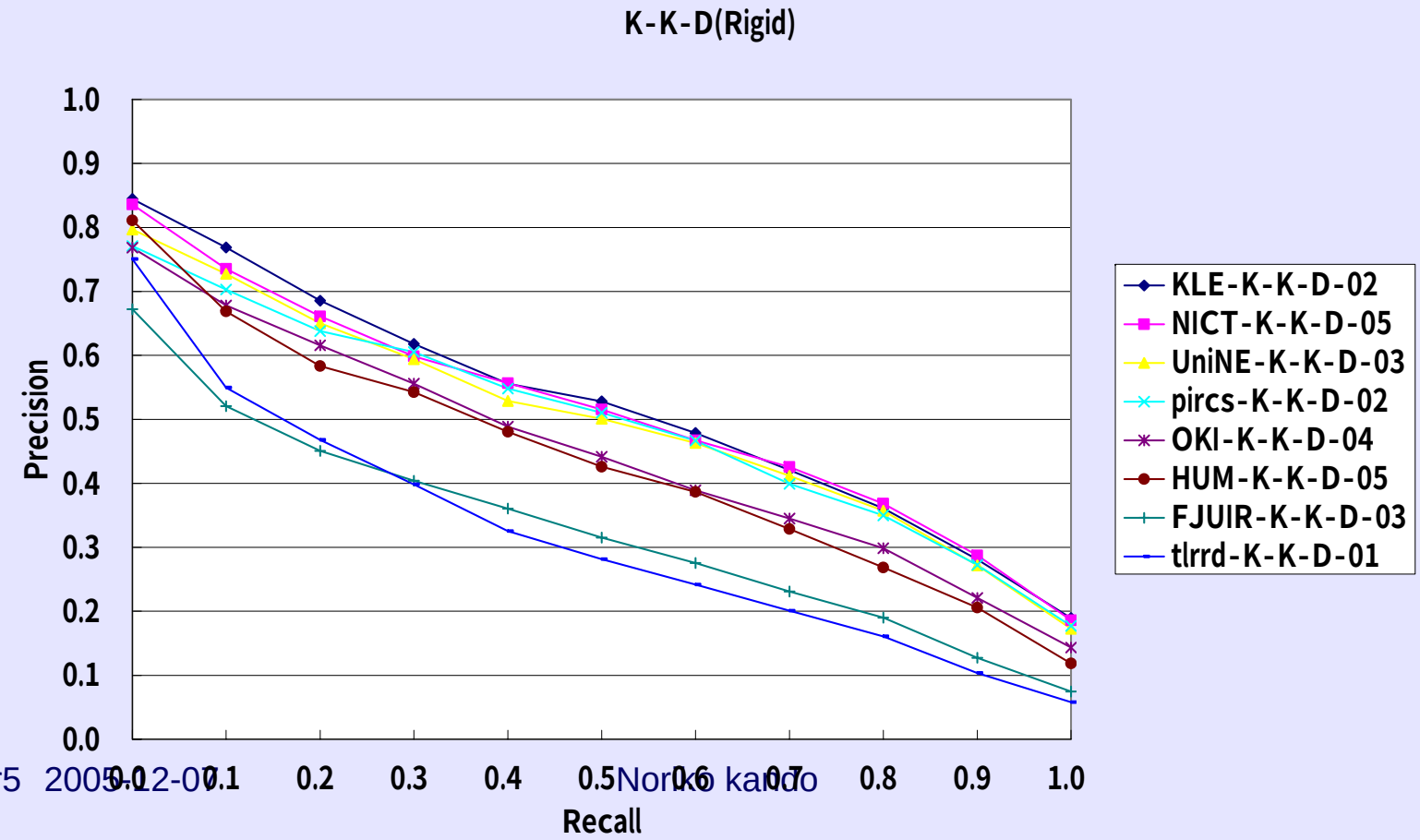
Evaluation (3)

- SLIR: J-J-D (Rigid) – top 8 groups



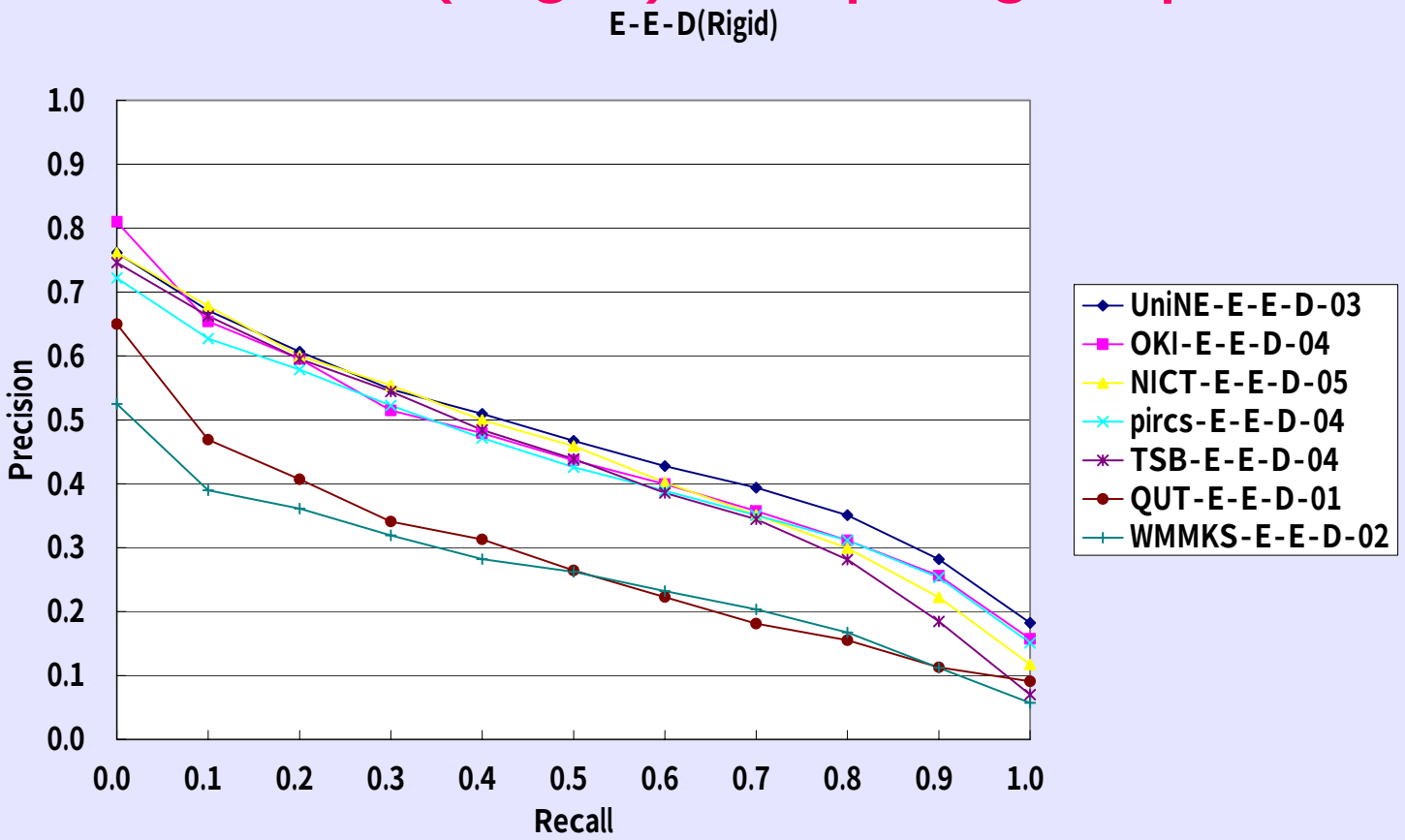
Evaluation (4)

- SLIR: K-K-D (Rigid) – top 8 groups



Evaluation (5)

- SLIR: E-E-D (Rigid) – top 7 groups



Evaluation (6)

- BLIR – Comparison of MAP values between best SLIR and best BLIR runs (D-run, Rigid)

	C-C .4826			J-J .3823	
J-C	.1568	32.5%	C-J	.2471	64.6%
K-C	.0377	7.8%	K-J	.2799	73.2%
E-C	.2682	55.6%	E-J	.2981	78.0%
	K-K .5079			E-E .4581	
C-K	.3263	64.2%	C-E	.4042	88.2%
J-K	.4511	88.8%	J-E	.4135	90.3%
E-K	.4092	80.6%	K-E	.1003	21.9%

Evaluation (7)

- MLIR – Best runs

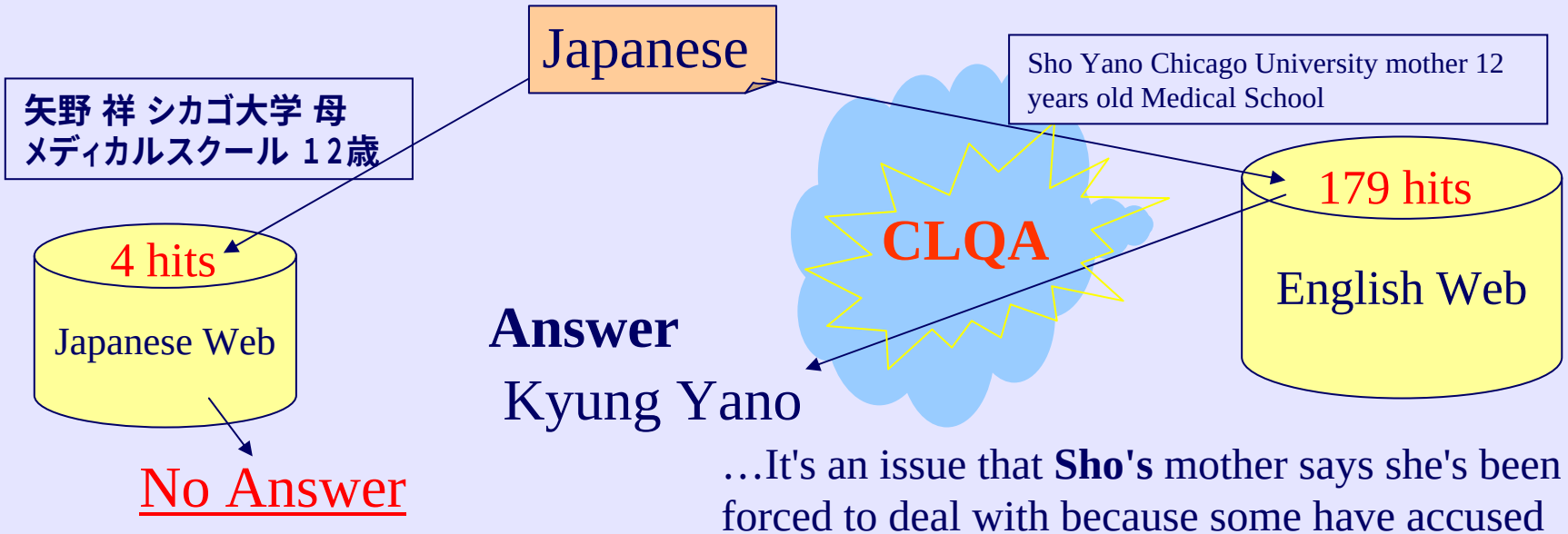
Run-type	MAP
C-CJKE	.2052
J-CJKE	.1890
K-CJKE	.1347
E-CJKE	.2695

Cross-Language Question Answering

Task Organizers
Kuang-hua Chen, NTU
Chuan-Jie Lin , Nat Taiwan Ocean U
Yutaka Sakaki, ATR

Necessity for Cross-Lingual QA

Question 「12歳でシカゴ大学メディカルスクールに入学した矢野祥君のお母さんの名前は？」
"What is mother's name of the student who goes to the University of Chicago Medical School at 12 years old."

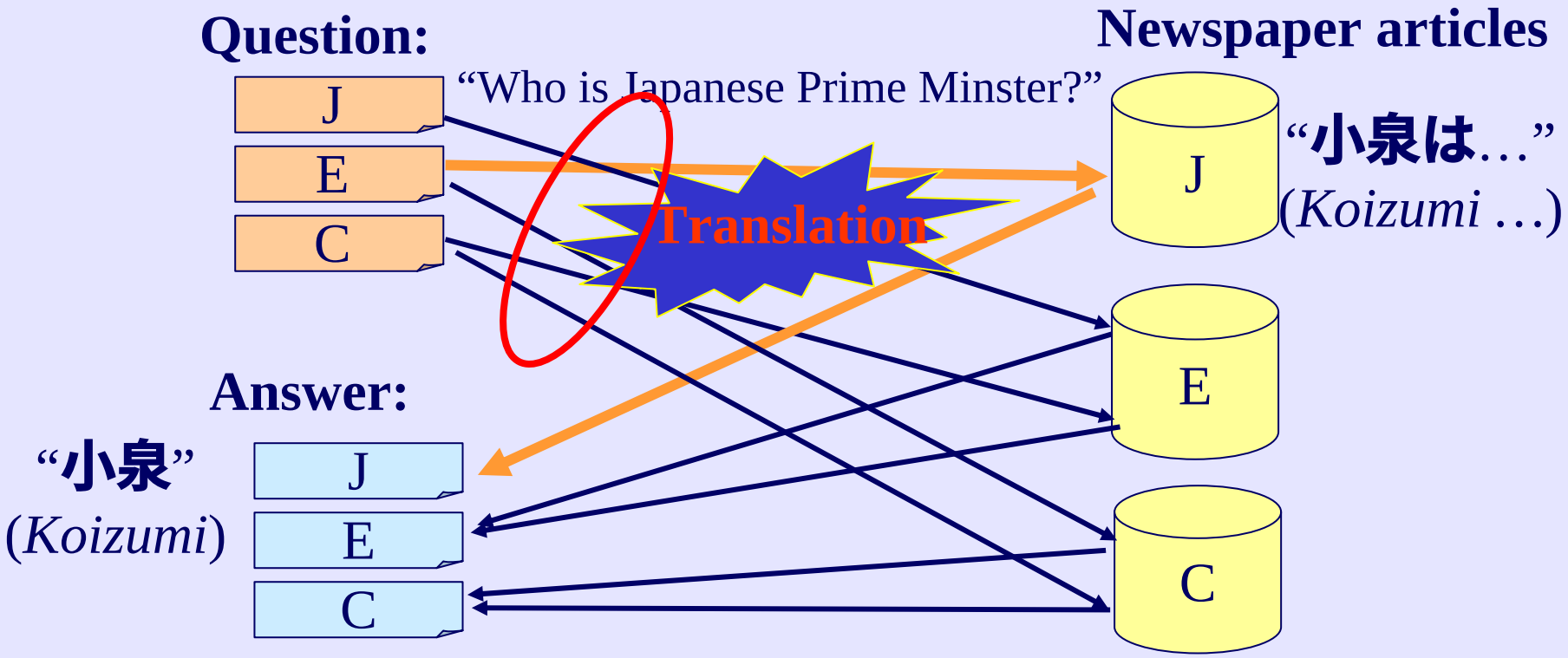


Answer
Kyung Yano

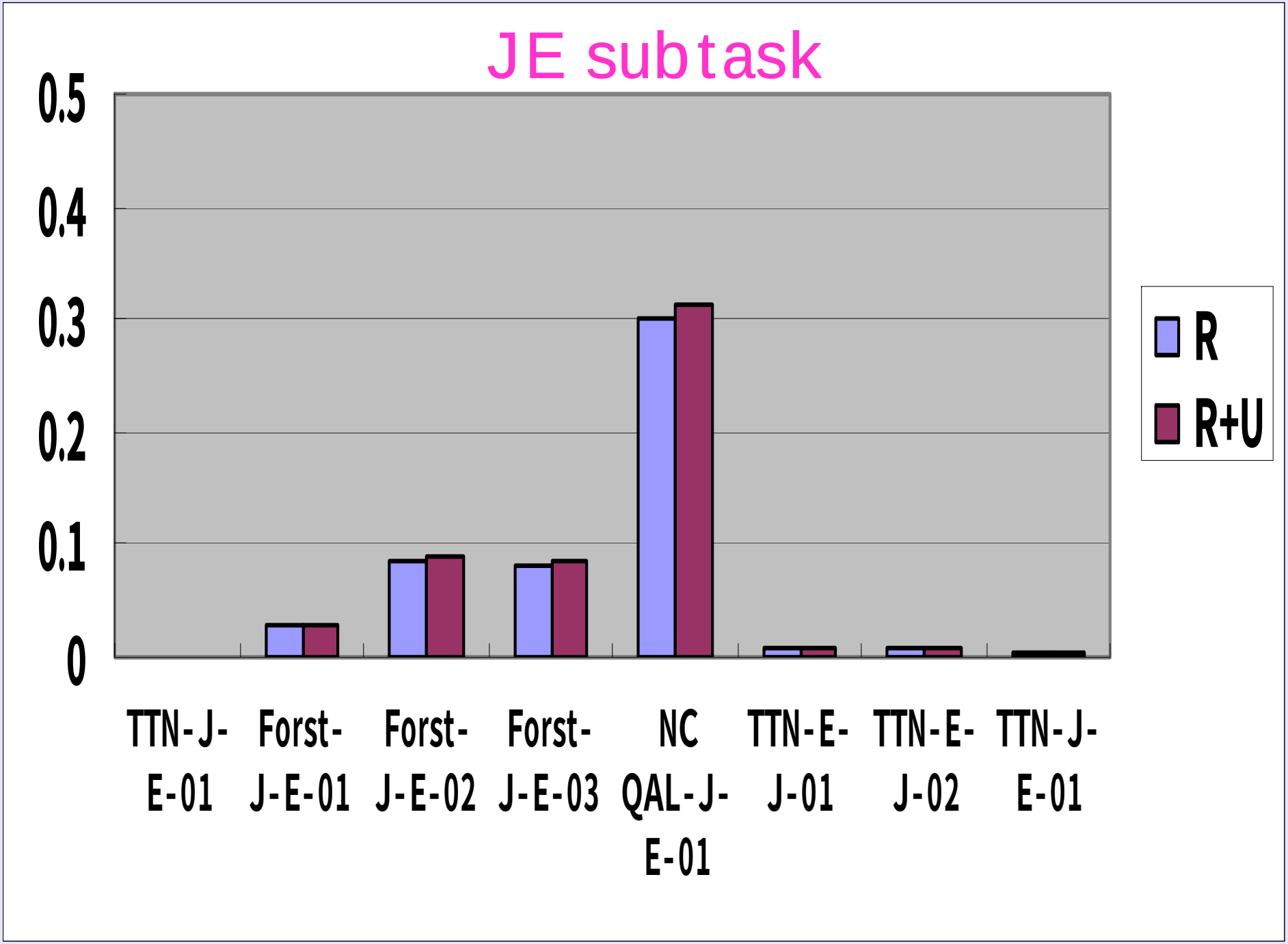
…It's an issue that **Sho's** mother says she's been forced to deal with because some have accused her of pushing her son too far too fast. "I am the mother of this child," says **Kyung Yano**. …

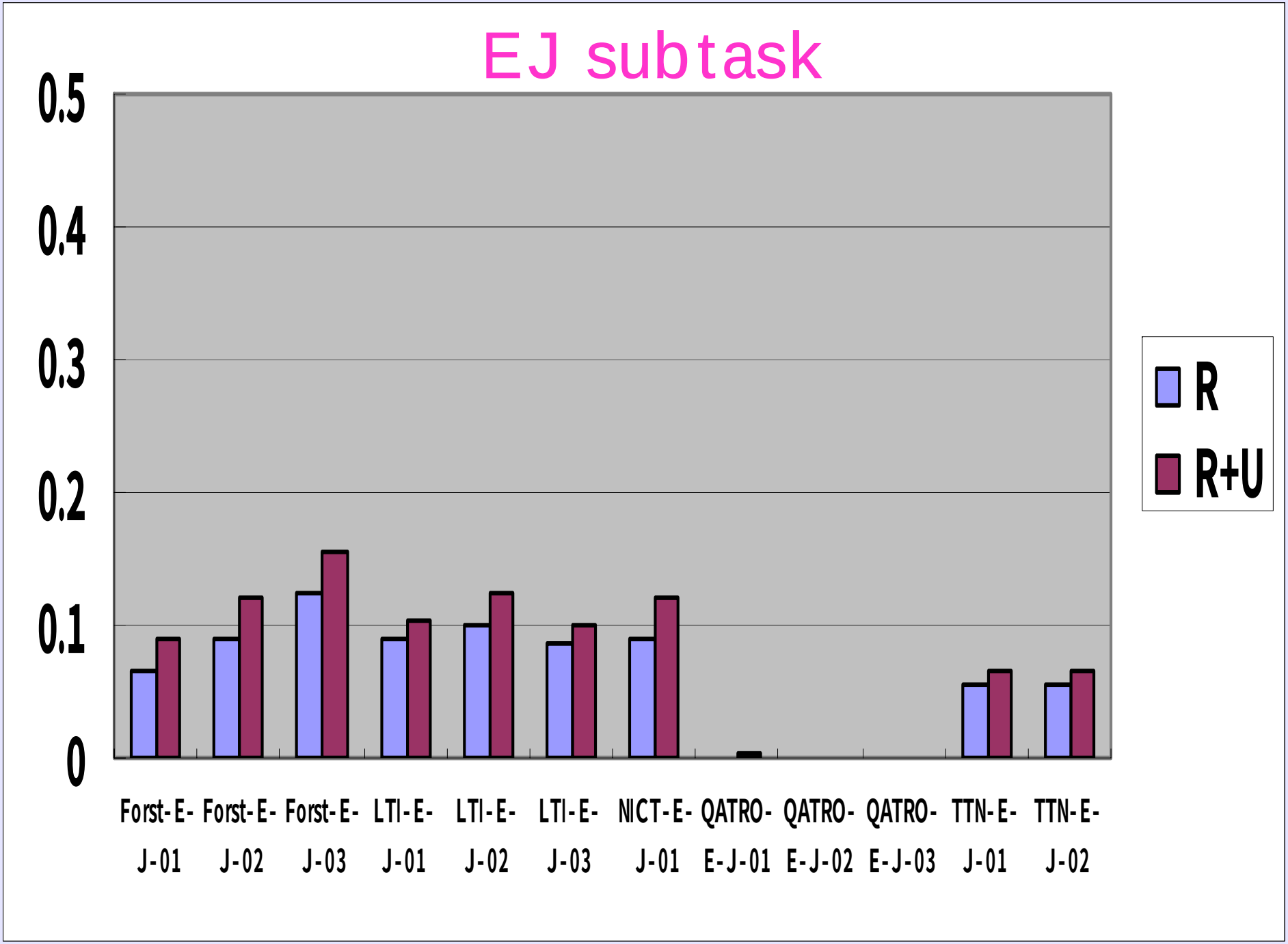
NTCIR Cross-Lingual Question Answering (CLQA1)

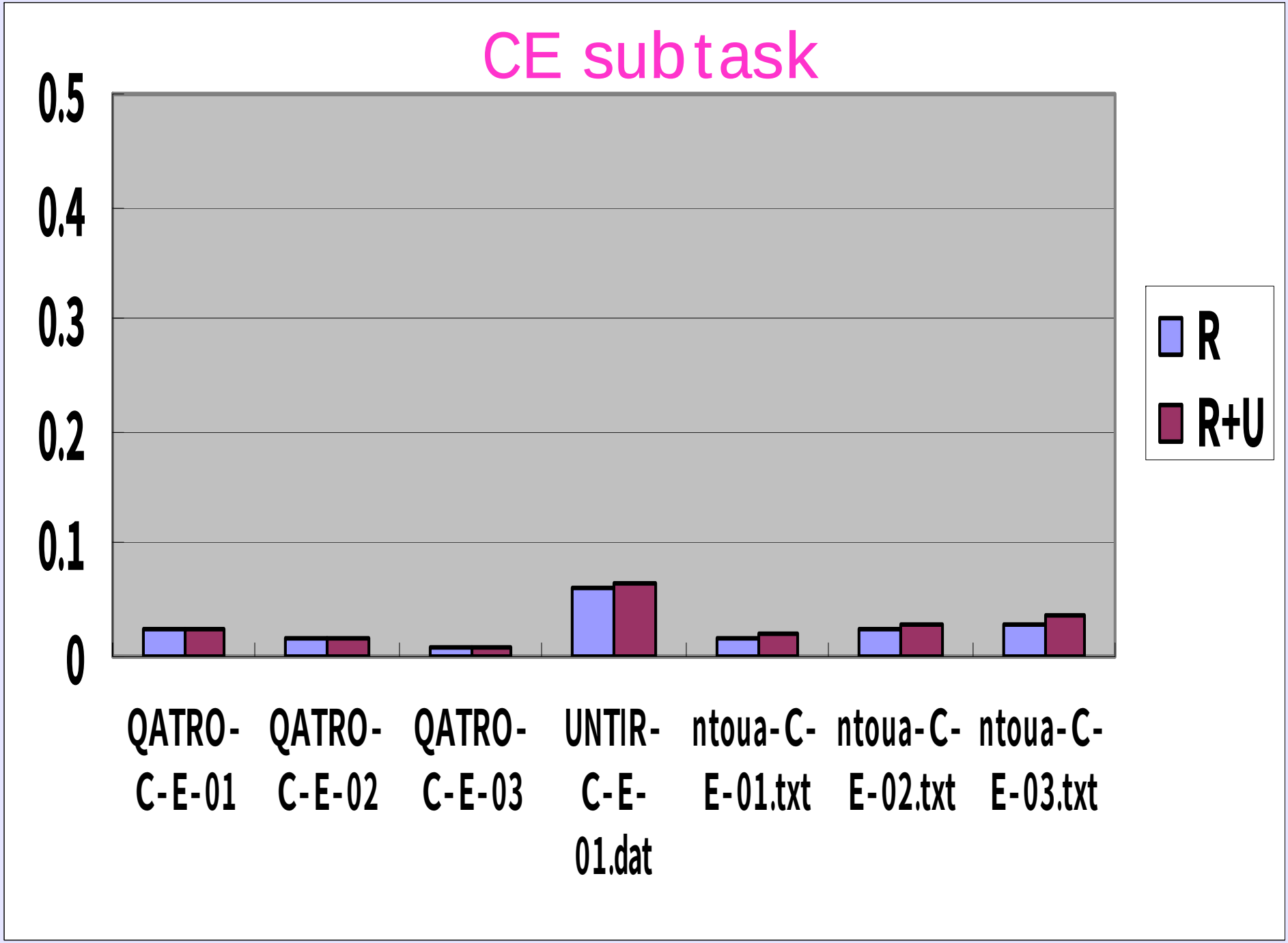
Language pairs: J->E, E->J, C->E, C->C, E->C
Target: Questions about named entities (PERSON, DATE, SPEED ..



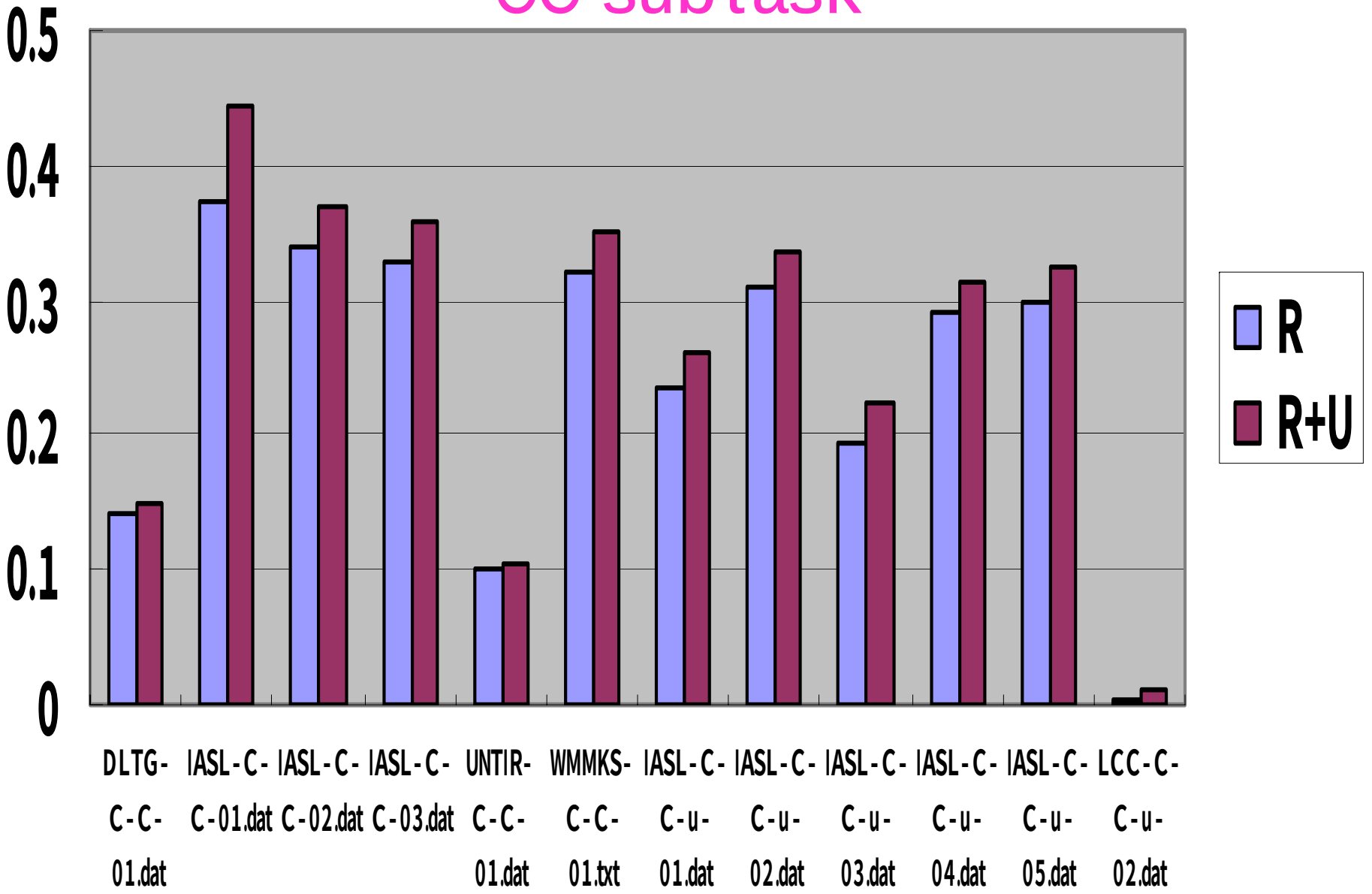
Organizers:
Japanese: Yutaka Sasaki (ATR)
Chinese: Hsin-Hsi Chen
Kuang-hua Chen
Chuan-Jie Lin (NTU)

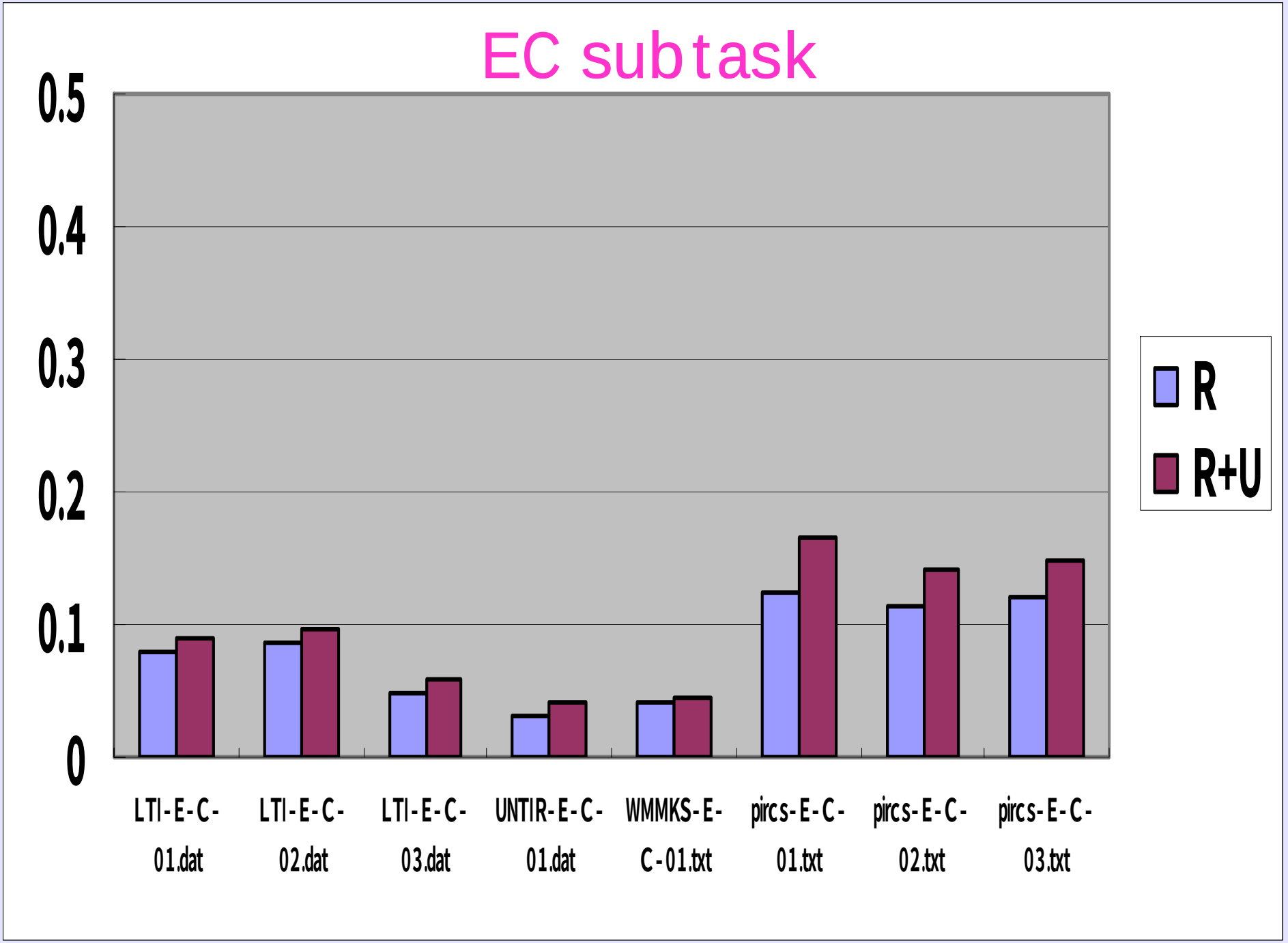

JE subtask
0.5
0.4
0.3
0.2
0.1
0
TTN-J-E-01
Forst-J-E-01
Forst-J-E-02
Forst-J-E-03
NC QAL-J-E-01
TTN-E-J-01
TTN-E-J-02
TTN-J-E-01
R
R+U

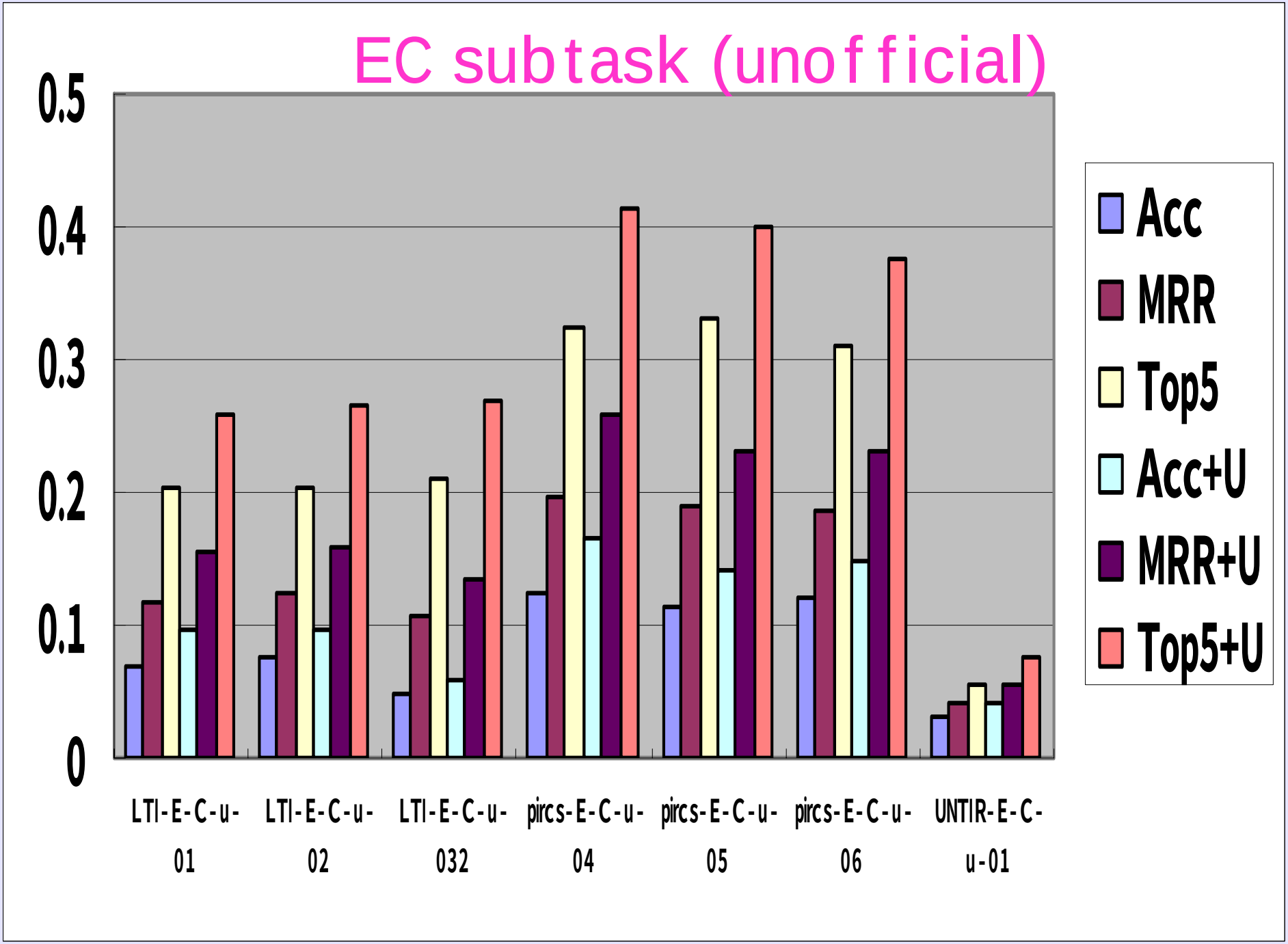

EJ subtask
0.5
0.4
0.3
0.2
0.1
0
Forst-E-J-01
Forst-E-J-02
Forst-E-J-03
LTI-E-J-01
LTI-E-J-02
LTI-E-J-03
NICT-E-J-01
QATRO-E-J-01
QATRO-E-J-02
QATRO-E-J-03
TTN-E-J-01
TTN-E-J-02
R
R+U


CE subtask
0.5
0.4
0.3
0.2
0.1
0
R
R+U
QATRO-C-E-01
QATRO-C-E-02
QATRO-C-E-03
UNTIR-C-E-01.dat
ntoua-C-E-01.txt
ntoua-C-E-02.txt
ntoua-C-E-03.txt

CC subtask


0.5
0.4
0.3
0.2
0.1
0
R
R+U
DLTG-C-C-01.dat
IASL-C-C-01.dat
IASL-C-C-02.dat
IASL-C-C-03.dat
UNTIR-C-C-01.dat
WMMKS-C-C-01.txt
IASL-C-C-u-01.dat
IASL-C-C-u-02.dat
IASL-C-C-u-03.dat
IASL-C-C-u-04.dat
IASL-C-C-u-05.dat
LCC-C-C-u-02.dat


EC subtask
0.5
0.4
0.3
0.2
0.1
0
LTI-E-C-01.dat
LTI-E-C-02.dat
LTI-E-C-03.dat
UNTIR-E-C-01.dat
WMMKS-E-C-01.txt
pircs-E-C-01.txt
pircs-E-C-02.txt
pircs-E-C-03.txt
R
R+U


EC subtask (unofficial)
0.5
0.4
0.3
0.2
0.1
0
LTI-E-C-u-01
LTI-E-C-u-02
LTI-E-C-u-032
pircs-E-C-u-04
pircs-E-C-u-05
pircs-E-C-u-06
UNTIR-E-C-u-01
Acc
MRR
Top5
Acc+U
MRR+U
Top5+U

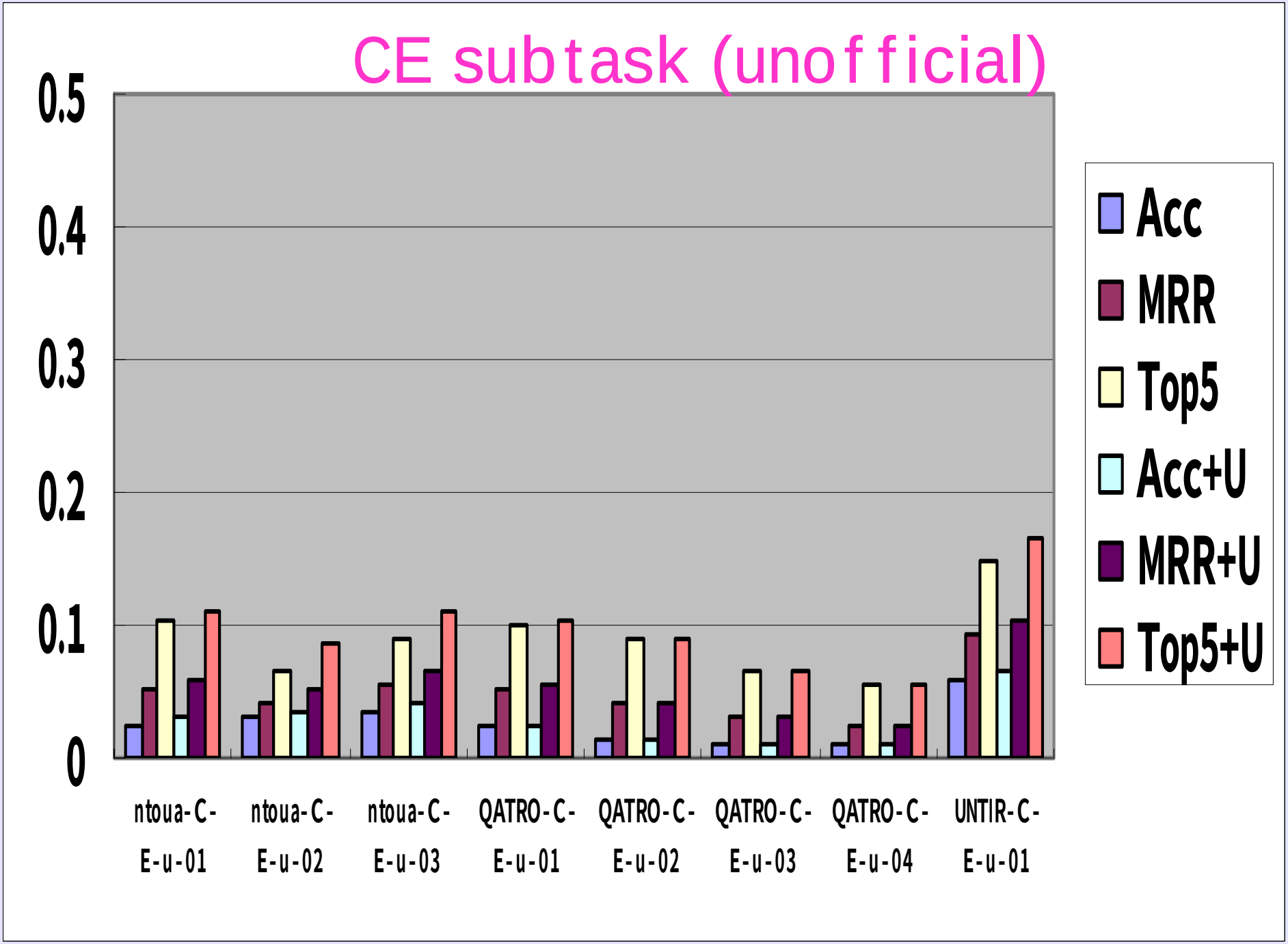

CE subtask (unofficial)
0.5
0.4
0.3
0.2
0.1
0
ntoua-C-E-u-01
ntoua-C-E-u-02
ntoua-C-E-u-03
QATRO-C-E-u-01
QATRO-C-E-u-02
QATRO-C-E-u-03
QATRO-C-E-u-04
UNTIR-C-E-u-01
Acc
MRR
Top5
Acc+U
MRR+U
Top5+U

Table 4. Monolingual vs. Cross-lingual

Run ID	# Correct		Accuracy	
	# R	# R+U	R	R+U
Forst-E-J-03	25	31	12.5	15.5
Forst-J-J	34	53	17.0	26.5
			+4.5	+11.0
LTI-E-J-02	20	25	10.0	12.5
LTI-J-J	16	40	8.0	20.0
			-2.0	+7.5
NCQAL-J-E-01	60	63	30.0	31.5
NCQAL-E-E	74	85	37.0	42.5
			+7.0	+11.0
NICT-E-J-01	18	24	9.0	12.0
NICT-J-J	34	53	17.0	26.5
			+8.0	+14.5
QATRO-E-J-01*	0	1	0.0	0.5
QATRO-J-J	4	9	2.0	4.5
			+2.0	+4.0
QATRO-J-E-01*	2	2	1.0	1.0
QATRO-E-E	11	16	5.5	8.0
			+4.5	+7.0
TTN-E-J-01	11	13	5.5	6.5
TTN-J-J	22	35	11.0	17.5
			+5.5	+11.0

Table 6. EJ Subtask

Run ID	# Correct		Accuracy	
	# R	# R+U	R	R+U
LTI-E-J-01	18	21	9.0	10.5
LTI-E-J-02	20	25	10.0	12.5
LTI-E-J-03	17	20	8.5	10.0
Forst-E-J-01	13	18	6.5	9.0
Forst-E-J-02	18	28	9.0	14.0
Forst-E-J-03	25	31	12.5	15.5
NICT-E-J-01	18	24	9.0	12.0
QATRO-E-J-01*	0	1	0.0	0.5
QATRO-E-J-02*	0	0	0.0	0.0
QATRO-E-J-03*	0	0	0.0	0.0
TTN-E-J-01	11	13	5.5	6.5
TTN-E-J-02	11	13	5.5	6.5

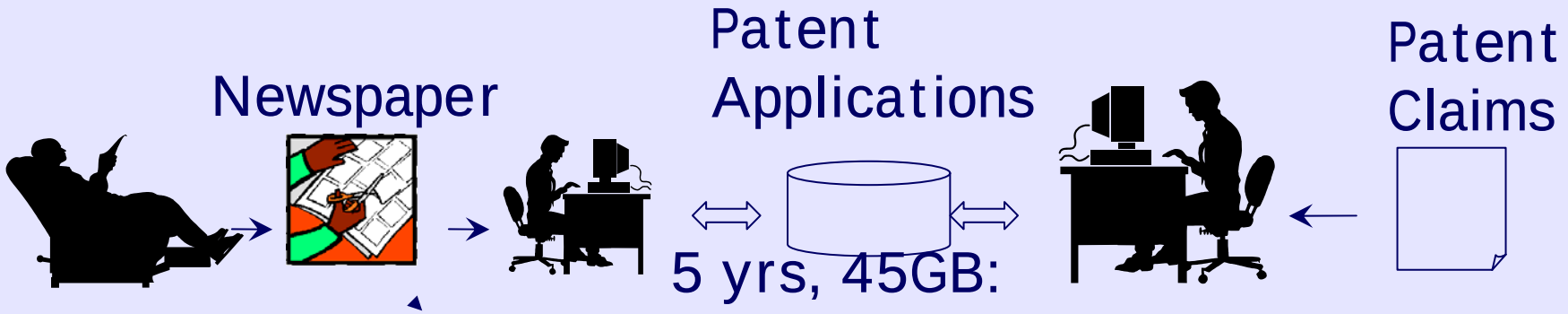
Table 8. CC Subtask

Run ID	# Correct		Accuracy	
	# R	# R+U	R	R+U
DLTG-C-C-01	28	30	14.0	15.0
IASL-C-C-01	75	89	37.5	44.5
IASL-C-C-02	68	74	34.0	37.0
IASL-C-C-03	66	72	33.0	36.0
lcc-C-C-01	20	21	10.0	10.5
UNTIR-C-C-01	20	21	10.0	10.5
WMMKS-C-C-01	64	70	32.0	35.0

Patent Retrieval Task

Task Organizers
Atsushi Fujii (Univ of Tsukuba)
Makoto Iwayama (TIT/Hitachi)
Noriko Kando (NII)

Patent Retrieval Tasks
situation & users' information seeking task



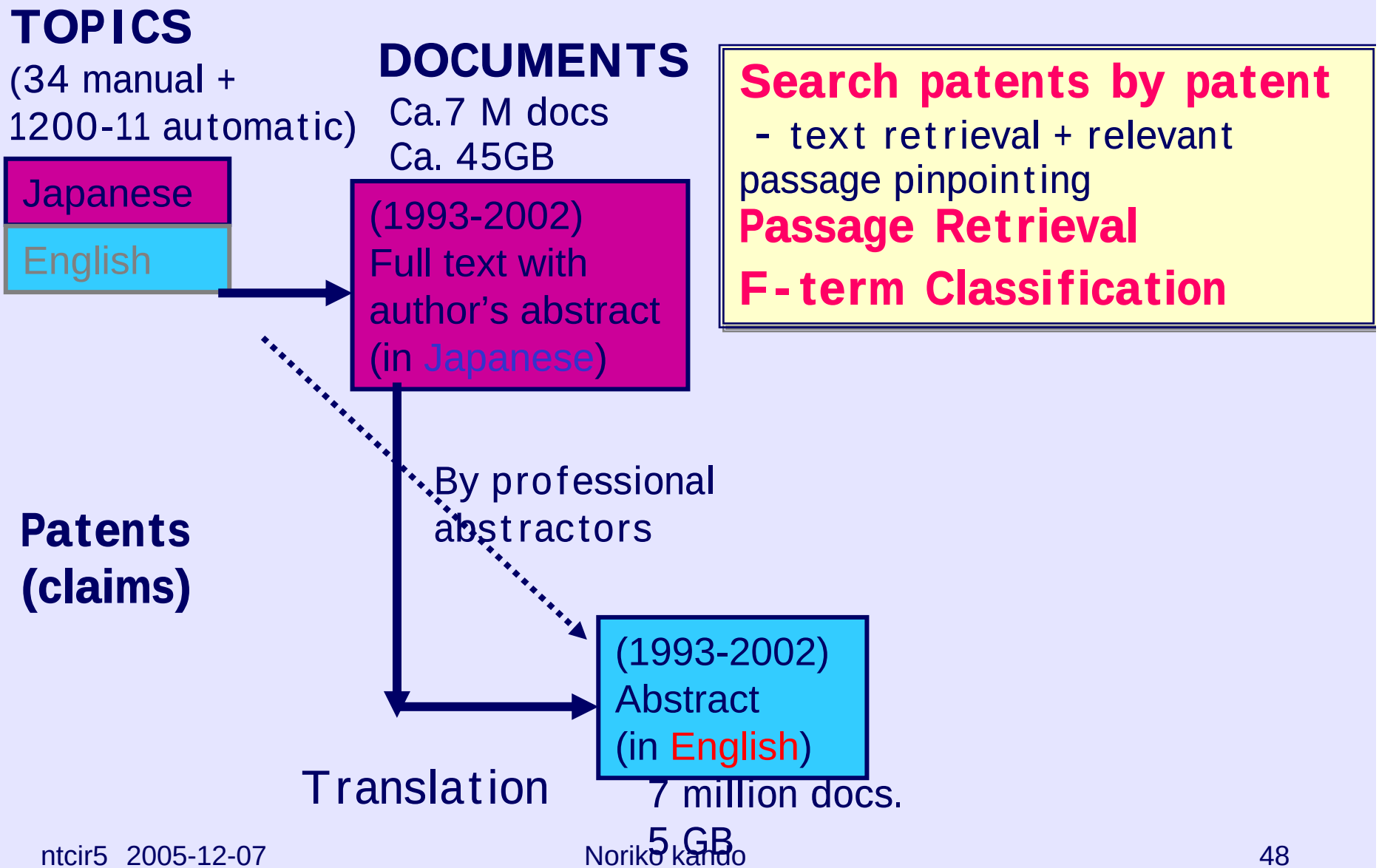
NTCIR-3 PATENT
(2001-2002)

Technological Survey:
Search patents by newspaper
End user: non-experts (ex. Business manager)

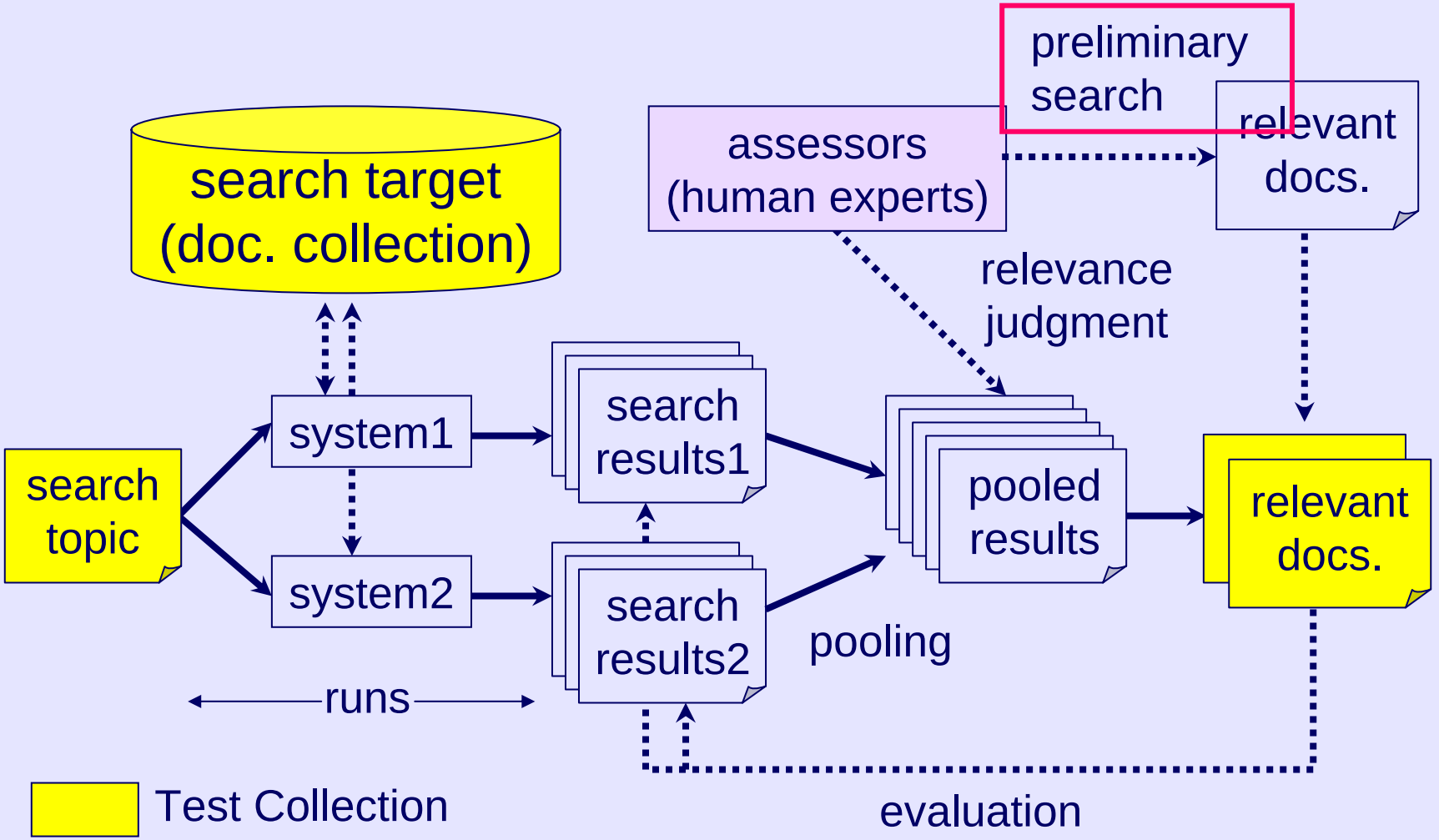
NTCIR-4,-5 PATENT
(2003-2004)(2004-2005)

From a claim of a new patent application, search patents that can invalidate the new patent application.
User: patent experts

NTCIR-5 Patent (2004-2005)



Test Collection Creation Procedure



Search topics

- Japanese patent application rejected by Japanese Patent Office (JPO)
- 34 main topics: selected and judged by human patent experts of “Japan Intellectual Property Association” (JIPA) (created at NTCIR-4)
- 1189 additional topics: applications rejected by JPO/ evaluate by using the citations only
- Quite few relevant documents

Example search topic

Date of filing

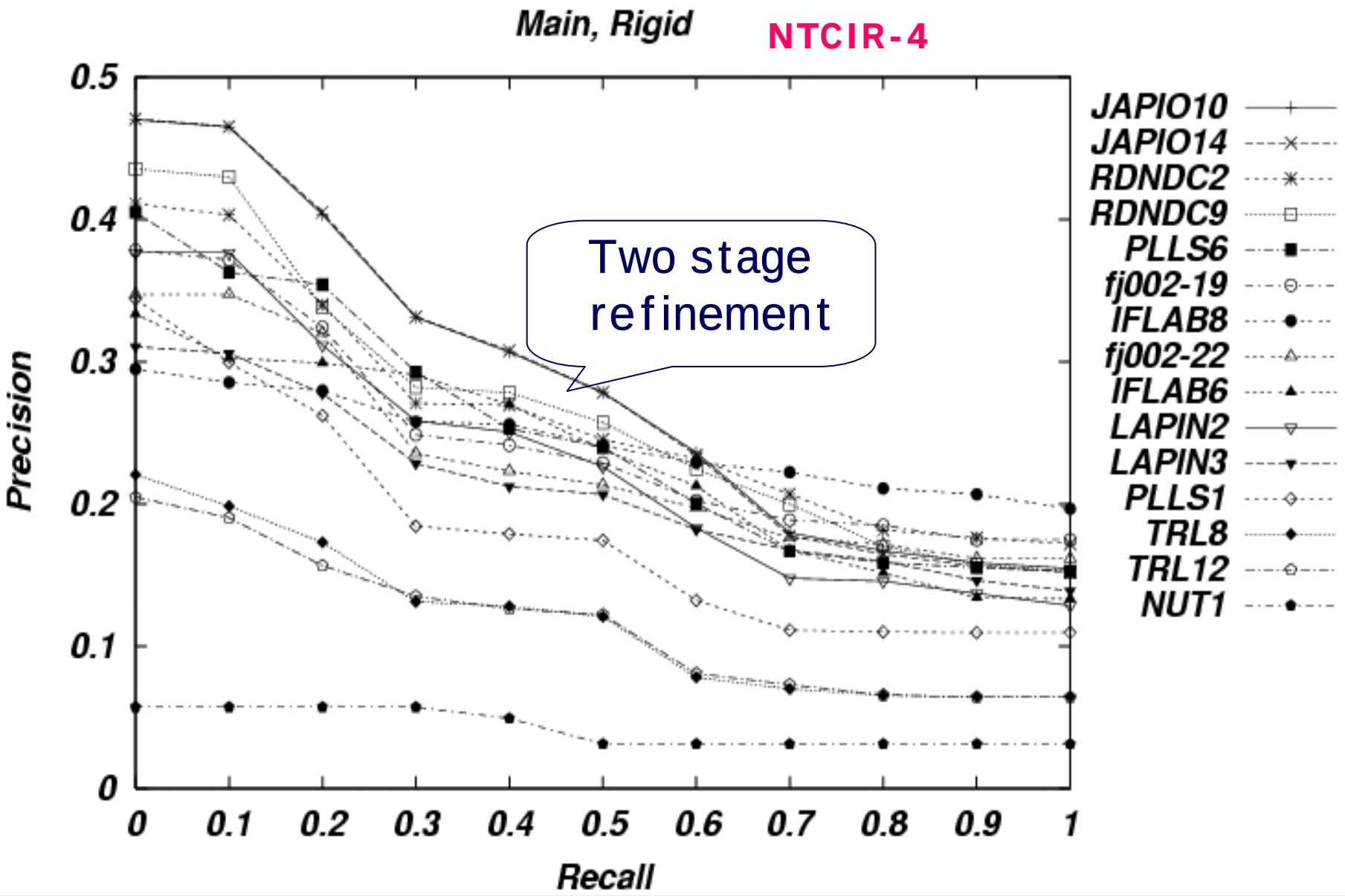
Relevant documents must be prior art, which had been open to the public before the topic patent was filed

<TOPIC>
<NUM>008</NUM>
<LANG>EN</LANG>
<FDATE>19960527</FDATE>
<CLAIM>(Claim 1) A sensor device, characterized in that an open recessed part is formed on a box-shaped forming base, a conductive film of a designated pattern is formed on the surface of the forming base including the inner surface of the recessed part, an element for a sensor is bonded to the recessed part,and the forming base is closed with a cover.</CLAIM>
...
</TOPIC>

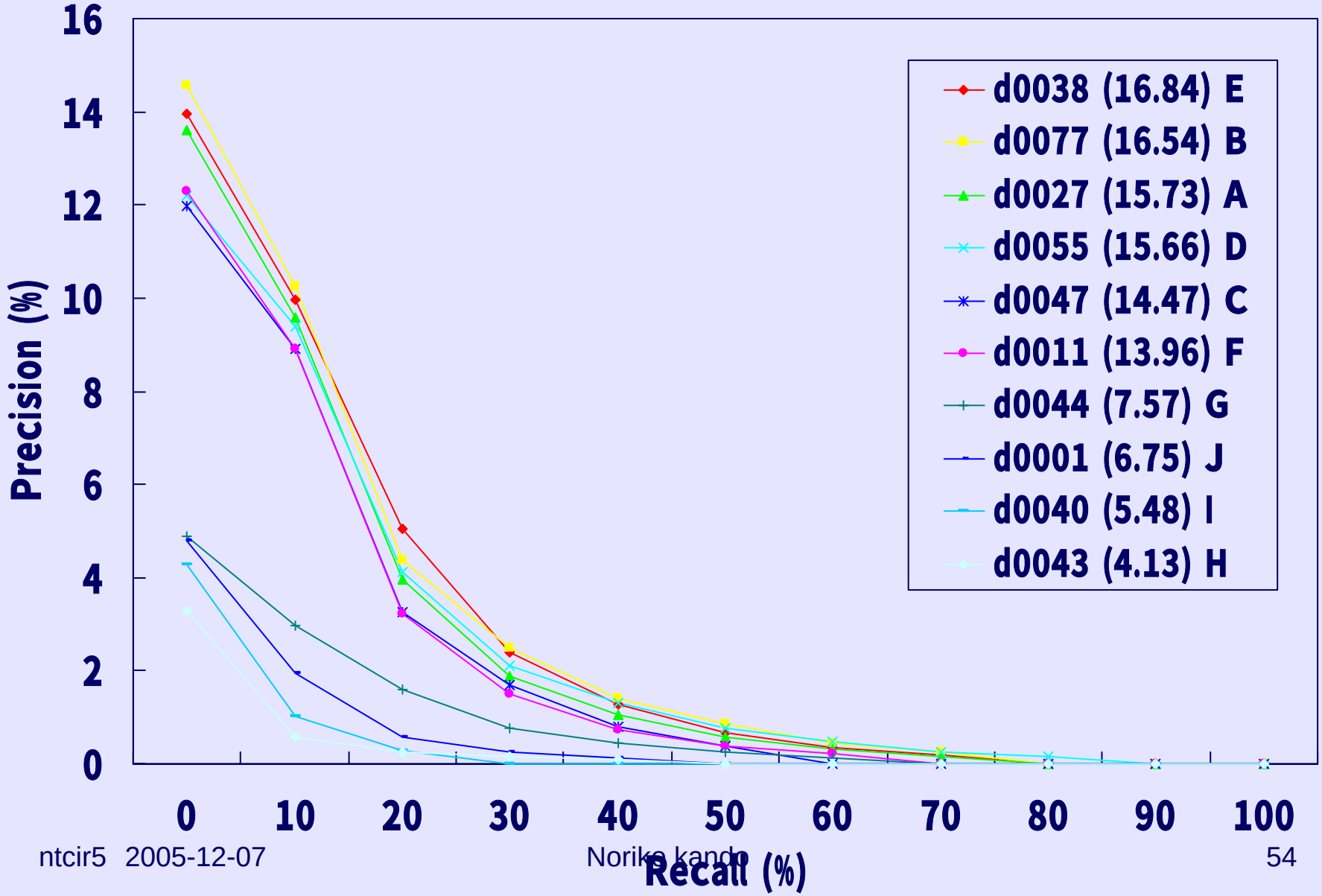
Target for invalidation

Relevance judgment

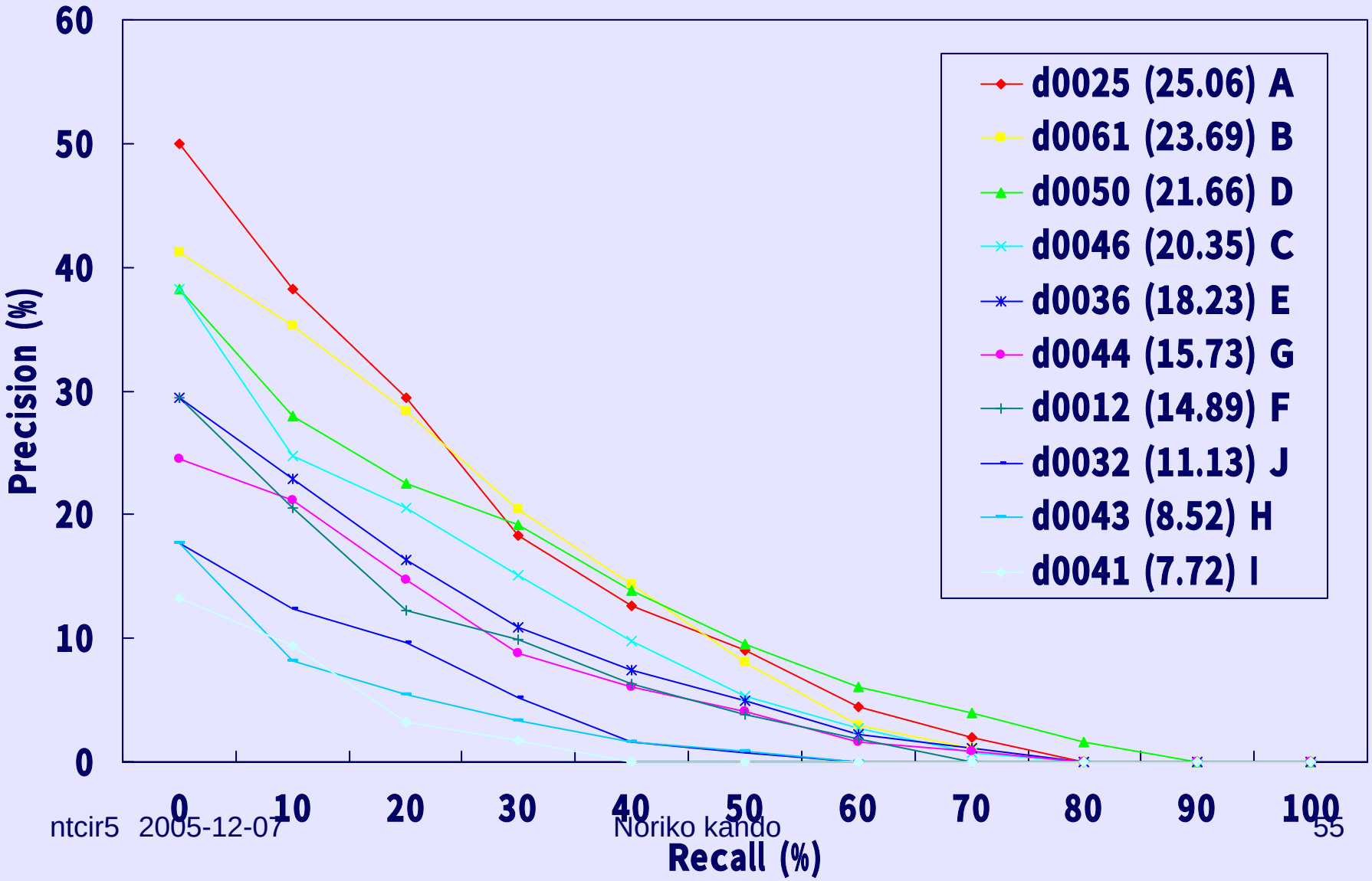
- Document-based relevant judgment
  - A: patent that can invalidate the topic claim
  - B: patent that can invalidate the topic claim, when used with other patents
- passage-based relevant judgment:
  - combinational relevance
- Submitted runs were evaluated by mean average precision (MAP)


Main, Rigid
NTCIR-4
Two stage refinement
Precision
Recall
0.5
0.4
0.3
0.2
0.1
0
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
JAPIO10
JAPIO14
RDNDC2
RDNDC9
PLLS6
fj002-19
IFLAB8
fj002-22
IFLAB6
LAPIN2
LAPIN3
PLLS1
TRL8
TRL12
NUT1

NTCIR-5 DocIR AB



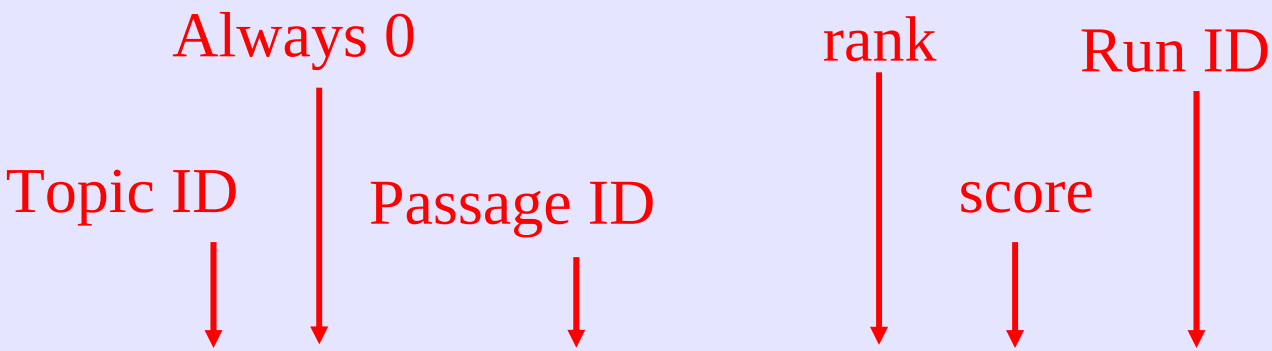
NTCIR-5 DocIR AB, Manual Queries



Passage Retrieval

- Provide Topics and Relevant documents
  - NTCIR-4 Topics 41
    - Dry runs 7, Formal runs 34
  - Relevant Docs 378
  - Sort the passages in the relevant docs

Ex. Results file passage retrieval



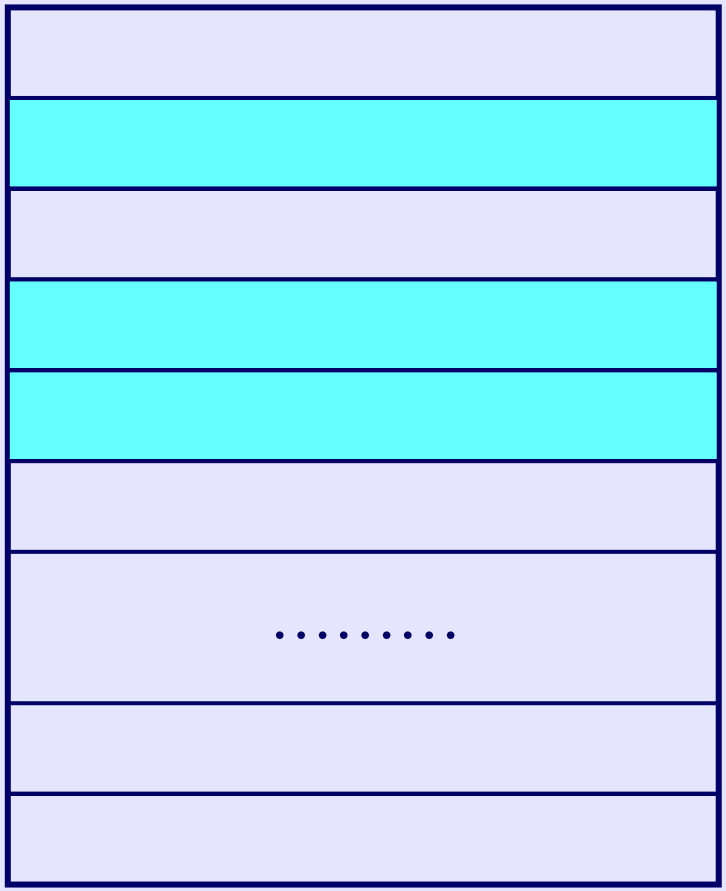
0001	0	1993-123456-5	1	9999	ntc1
0001	0	1993-123456-3	2	9999	ntc1
0001	0	1993-123456-0	3	9999	ntc1
0002	0	1994-000002-3	4	9999	ntc1
0002	0	1994-000002-1	5	9999	ntc1

...

Evaluation of passage Retrieval

- MAP
  - See both Recall and Precision
- Expected Search Length (ESL)
  - See Precision

Ex. evaluation by ESL



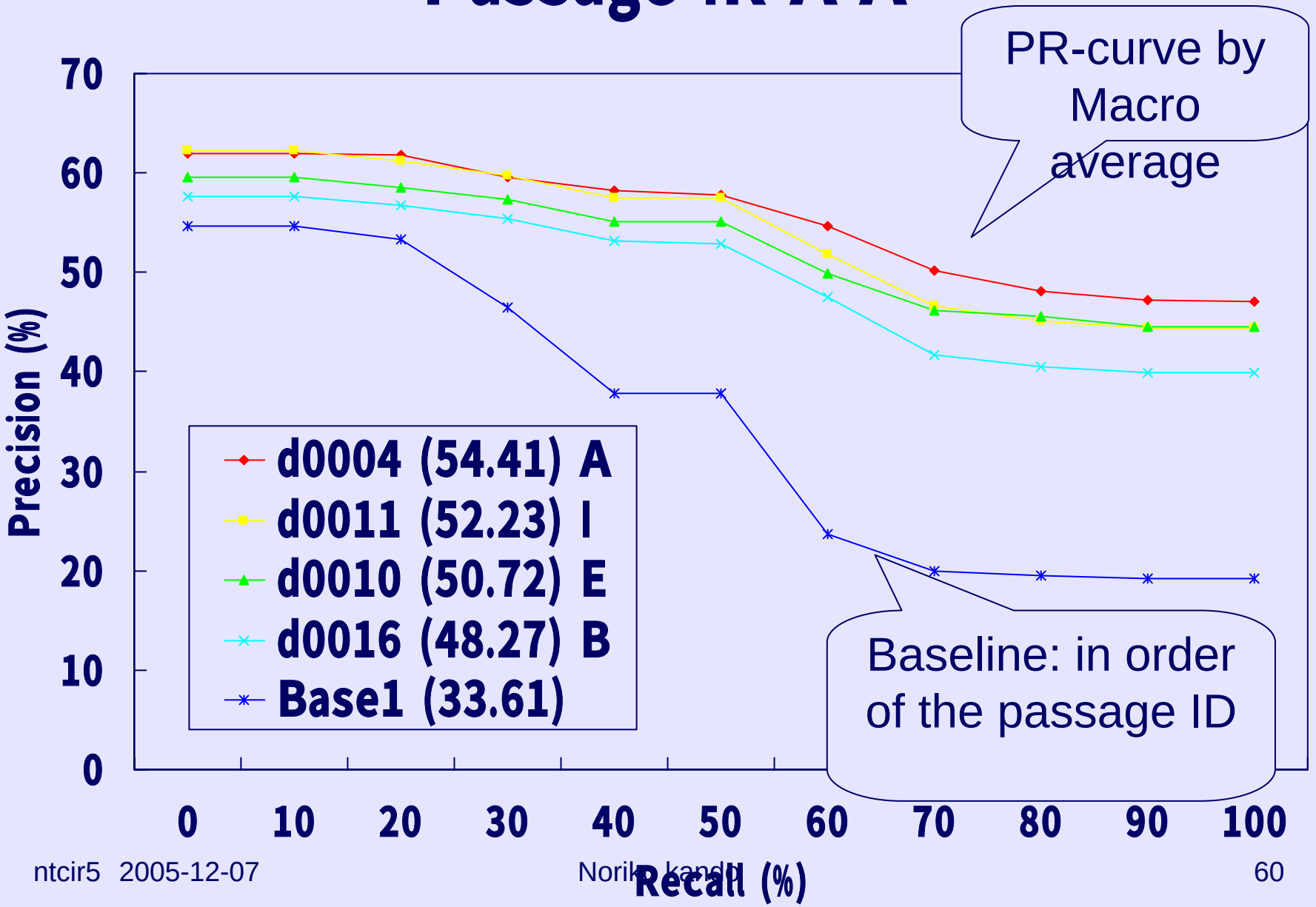
Search length = 5

Relevant passage

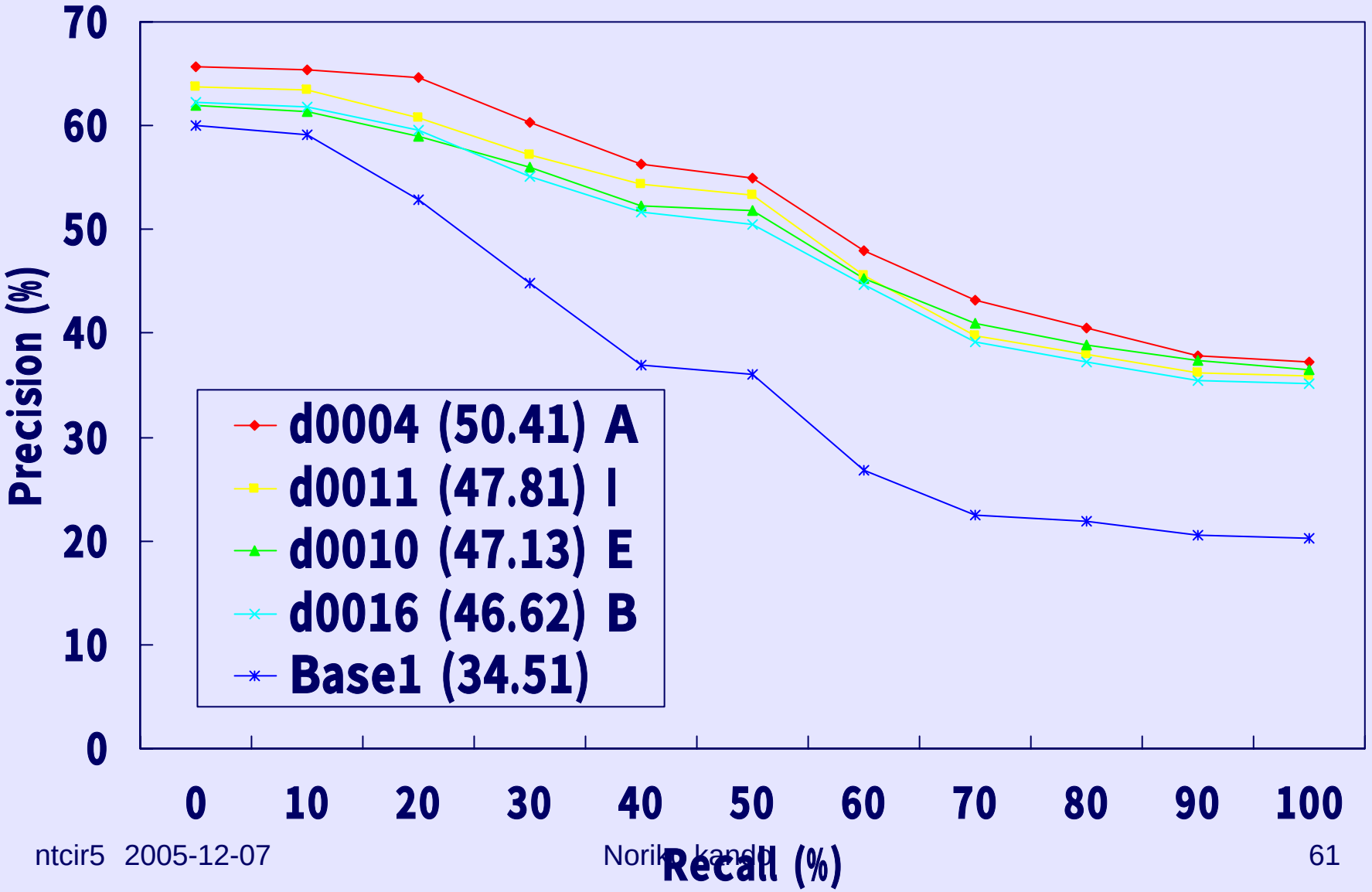
Relevant docs

- evaluate by the number of passages (search length) that the user read by he/she obtains sufficient evidences
- average the search length by each rel doc

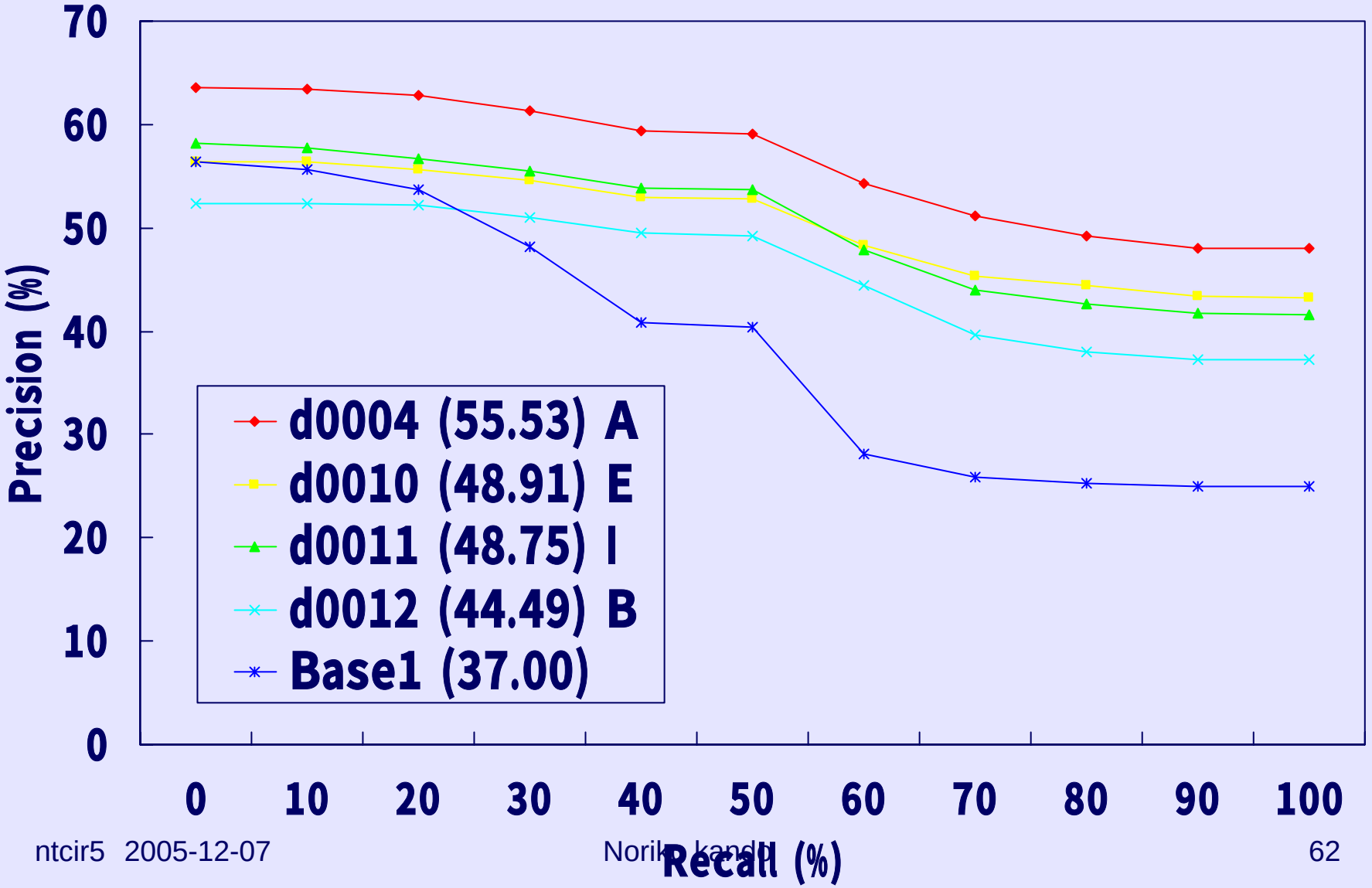
Passage IR A A



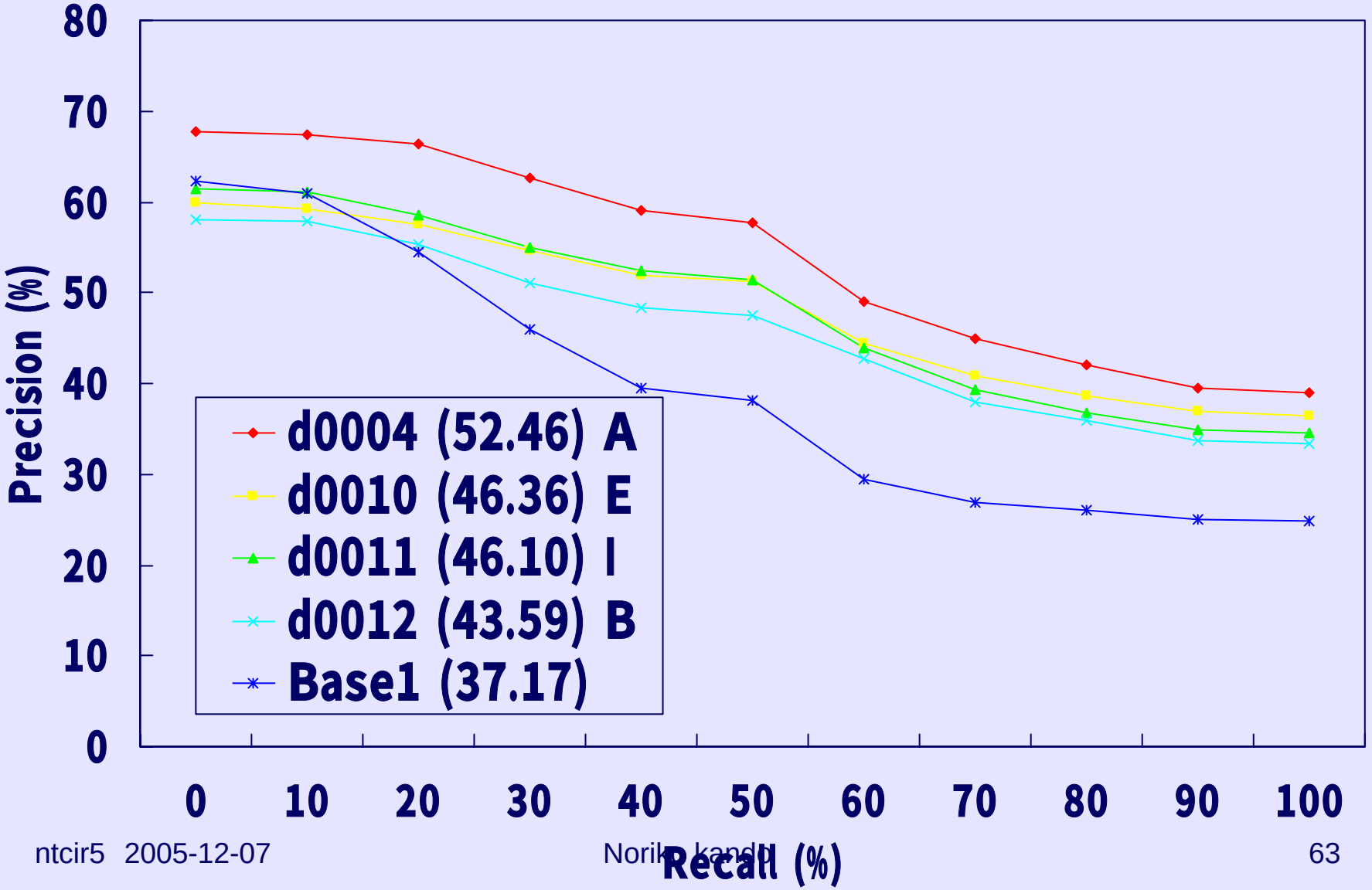
Passage IR A B



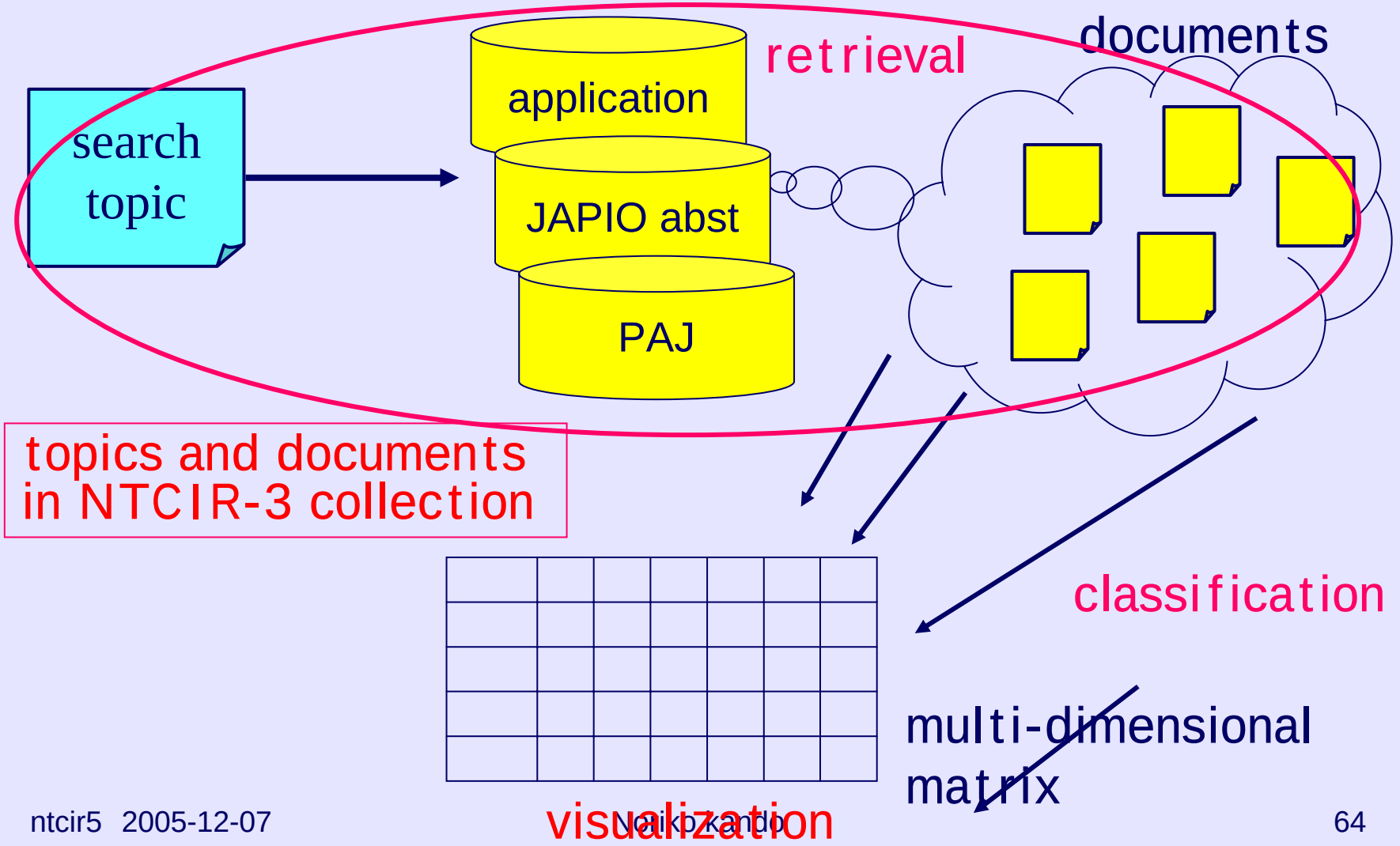
Passage IR B A



Passage IR B B



NTCIR-4 Feasibility Study: automatic patent map generation



Example (blue light-emitting diode)

given

problems to be solved

solutions

	crystalline	reliability	long operating life	emission stability	emission intensity
structure of active layer			1998-145000 1998-233554		
electrode composition		1998-107318		1998-190063 1998-209498	1998-209495
electrode arrangement		1998-215034 1998-223930	1998-242518	1998-173230 1998-209499 1998-256602	1998-242515 1998-270757
structure of light emitting element	1998-135516 1998-242586 1998-247761		1998-135514 1998-256668		1998-012923 1998-247745 1998-256597

participants identify lines and columns

NTCIR4 FS(patent map) Lesson learned

- Classification(Clustering) : very good
- Labeling the clusters: future work
- “Solution” only
- Too small # of topics
- Evaluation: insufficient
  - Can not cross system evaluation

NTCIR5 F term classification

- Use existing Classification (F terms)
  - Many topics
  - Cross-system Evaluation
- F term: multi-perspective classification
  - Can be used for Patent Map Automatic Creation

Tasks

- Topic classification
  - Provide Topic to each patent or Abstract
- F term classification
  - Provide F terms to patents (or abstracts) in a specific topic)

Purposes

- Topic classification：
  - Classification of the structured documents
- F term classification：
  - Multi-perspective classification

NTCIR5 QAC3

Tsuneaki Kato The University of Tokyo
Jun'ichi Fukumoto Ritsumeikan University
Fumito Masui Mie University

Information Access Dialogues

- Gathering Type
  - writing report on a specific topic
  - A (hidden) global topic and a series of related questions
  - Each consecutive question shares a local context
- Browsing Type
  - the topic of interest shifts through the dialogue
  - No global topic covers a whole dialogues
  - Each consecutive question shares a local context

Example of Series of Questions

- What genre does the "**Harry Potter**" series belong to?
- Who is the author?
- Who are the main characters in that series?
- When was the first volume published?
- What title does it have?
- How many volumes were published by 2001?
- How many languages has it been translated into?
- How many copies have been sold in Japan?

Series 02: Gathering Type

Example of Series of Questions

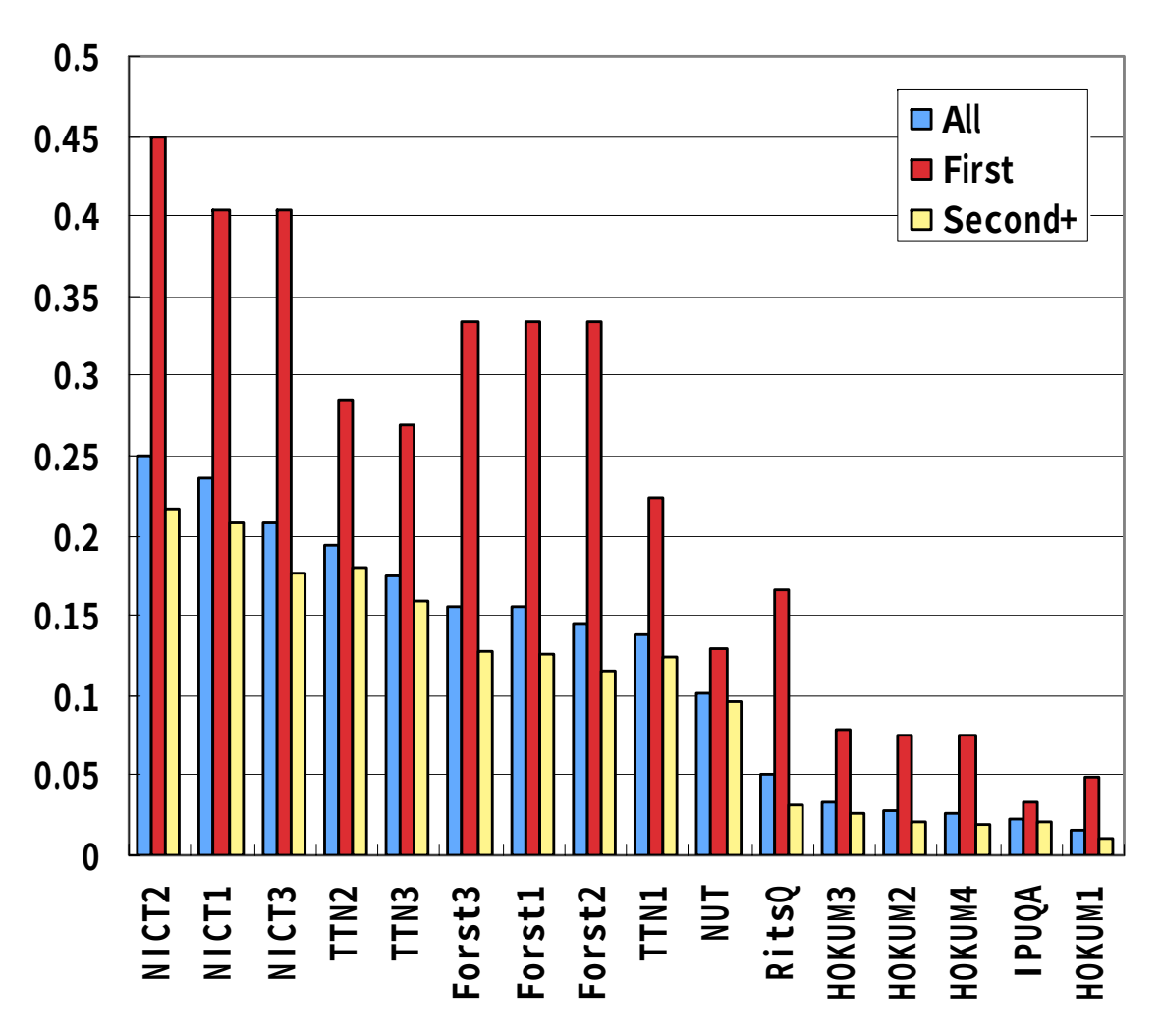
- Where was Universal Studio Japan constructed?
- Which train station is the nearest?
- Who is the actor who attended the ribbon-cutting ceremony on the opening day?
- What is the movie he was featured in that was released in the New Year season of 2001?
- What is the movie starring Kevin Costner released in the same season?
- What was the subject matter of that movie?
- What role did Costner play in that movie?

Series 24: Browsing Type

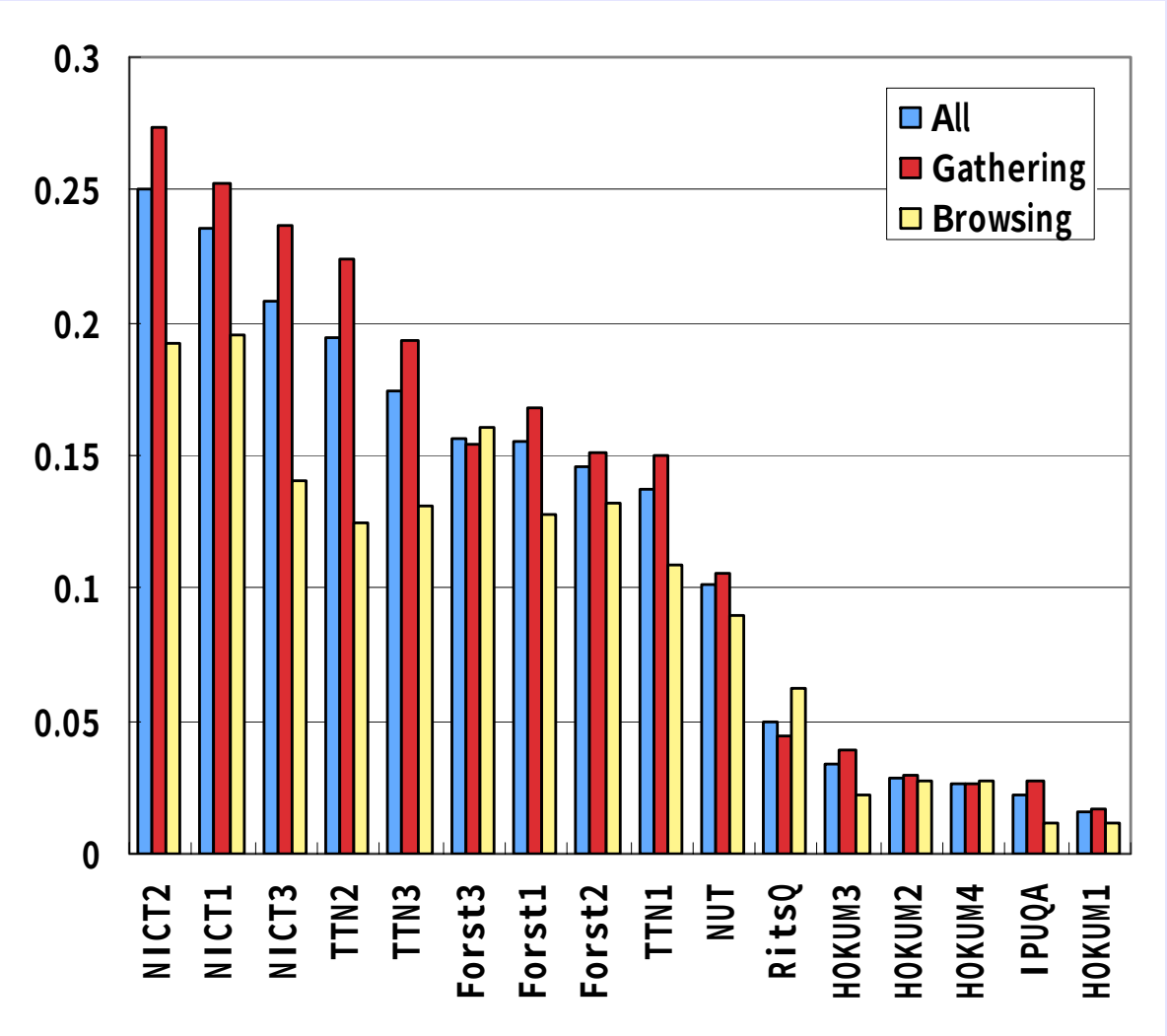
Design: Evaluation Measure

- An answer and a support (article)
- Mean Modified F measure,
  - For answer lists containing duplication denoting the same item
  - For questions with no answer
- The correctness of an answer is determined by human assessors within the given context
- 50 series(35 gathering and 15 browsing), 360 questions, 7.2 questions per series on average
- Reference Test Sets: anaphora resolved manually/automatically

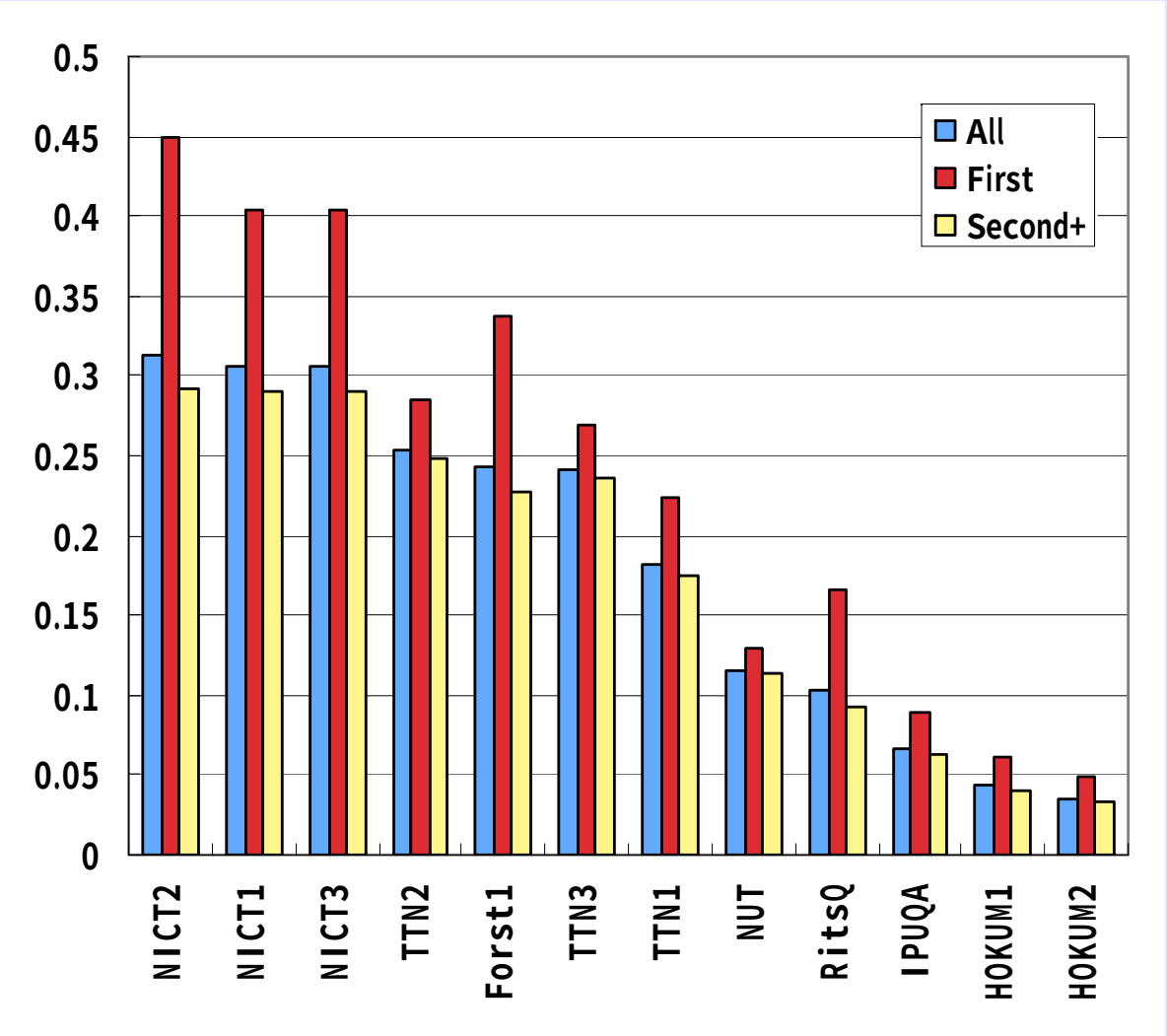
Evaluation by MMF



Differences on Series Type



Evaluation of Reference Run



Homework from NTCIR-4: Problems on Evaluation

One set of all the answers == F-measure

- Multiple answers and context

  Ex.

  - Q1-Countries in East Asia? Ans-PRC, ROC, N Korea, S Korea, UK
  - Q2-Capitals of these countries? Ans- Beijing, Taipei, Pyongyang, Soul, Tokyo

- Expression diversity and identification of the same answers

  Ex. A and B are the same or no[illegible] # of total correct answers and recall val[illegible]nds on such decision

- Major and minor answers

Wrong answer

Tokyo is not capital of UK. Correct answer for Q2 but this system produced wron answer for Q1.

WEB

Keizo Oyama, NII
Masao Takaku, NII
Masaharu Yoshioka, Hokkaido Univ

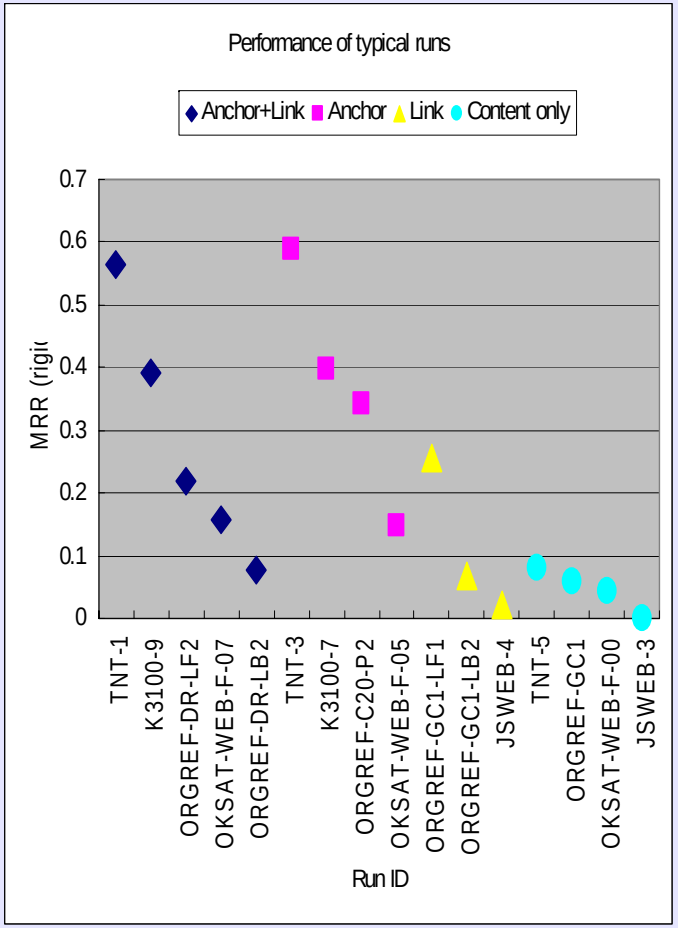
NTCIR-5 WEB

Navigational Retrieval Task 2

- **Goal: "Known Item Search"**.
  - To search for one or more representative Web pages on a known item.)
- **Topics:**
  - TITLE only is mandatory.
  - analyzing relationship among search techniques, topic types, search item categories and relevant page styles.
- **Submitted runs:** 35 (+28 by organizers) (+3 with trouble)
- **Relevance judgment:** *relevant, partially relevant, non-relevant*.
  "Representativeness" was judged based on every available information, e.g., provider of the page, content (text, images, etc.), URL, in/out-linked pages.

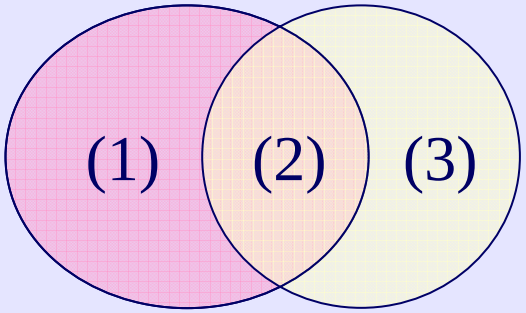
NTCIR-5 WEB Nav-2

- **Evaluation measures:** DCG and MRR at top-10 doc. level
- **Evaluation result:** Tendency on MRR & DCG
  - Several anchor-base systems performed best.
  - Link-base method or URL-base method made no contribution to anchor-base systems.
  - Several link-base systems performed fairly.
  - Content-base systems performed poorly.
- **Future work:**
  - considering document duplication
  - Verify stability of evaluation measures
  - Check comprehensiveness of assessment
  - Study on evaluation measures reflecting users' overall cost
  - topic-by-topic Analysis



Query Expansion Subtask

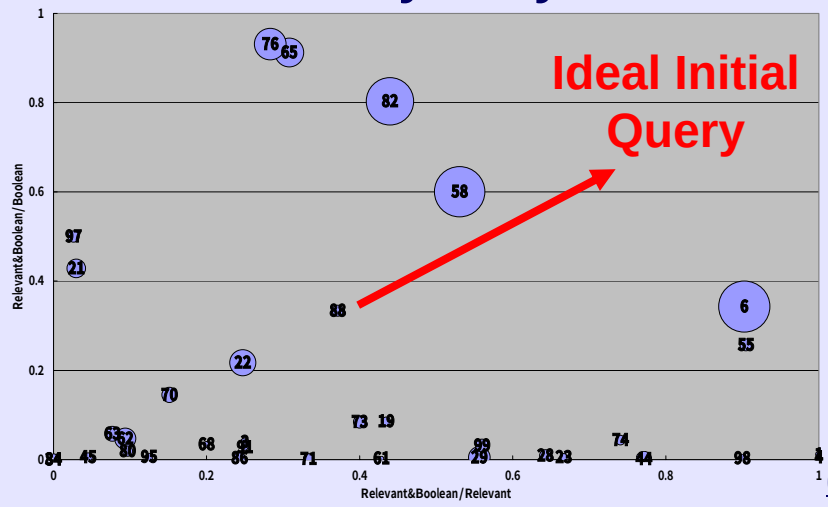
- Aim to establish framework for evaluating query term expansion technique.
- Enumerating bias for affecting query expansion results
  - Mismatch between initial query and relevant documents
  - Quality of relevant documents used by a system



Boolean query satisfied documents

Relevant documents

(2)
(1)+(2)



(2)
(2)+(3)

Ideal query: Both (1) and (3) are empty set

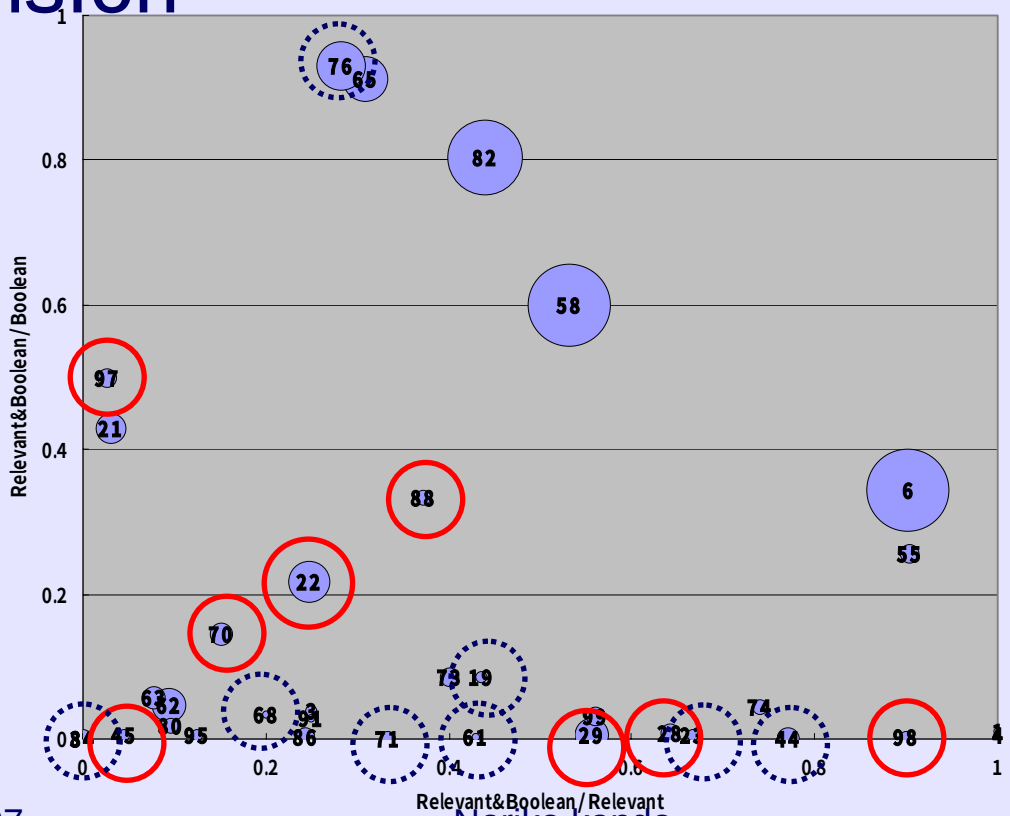
Large (1) = Ambiguous : needs context terms

Large (3) = Too strict : needs alternative terms

Characteristics of NTCIR-4 Web Survey Task Query (Title)

Topics where query term expansion works well

- No direct correlation between mismatch and effectiveness of the query term expansion

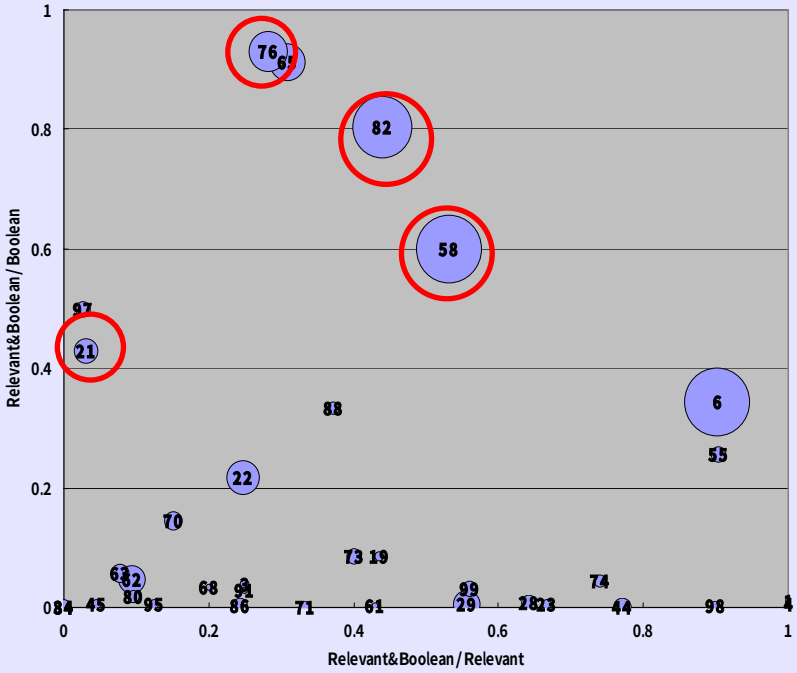


Topic where many number of performanc measures improve

Topic where small number of performanc measures improve

Effect of Feedback Documents Quality

- Comparison between automatic feedback and user feedback
  - Existence of topics whose retrieval results for user feedback is lower than automatic one
    - Run-Id: 0021, 0058, 0076, 0082
    - Those topics have higher R&B/B values compared with other topics
    - Good query expansion terms for these topics are terms that can be used as alternative terms for the initial query terms.
  - This result shows that non-relevant documents may be useful for finding alternative terms for initial query terms.



MuST: Multimodal Summarization for Trend Information

Tsuneaki Kato*1 Mitsunori Matsushita*2 N
Kando*3

*1 The University of Tokyo
*2 Nippon Telegraph and Telephone Corp.
*3 National Institute of Informatics

Characteristics

- To encouraging cooperative studies
  - Promoting discussion
  - Conforming communities
  - Constructing and accumulating resources
- No-evaluation
- Shared research resource
- Participants address common or related themes

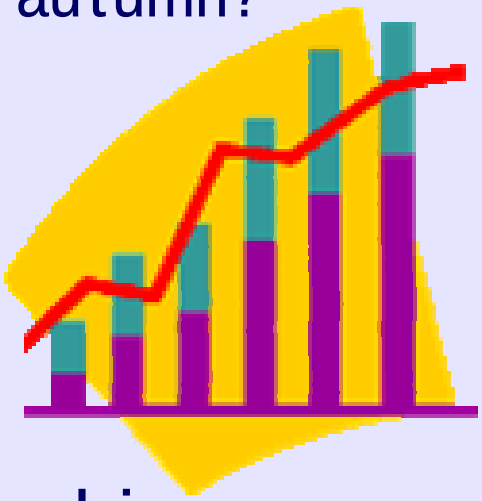
Multimodal summarization for Trend Information

Queries on trends

"How the price of gasoline shifted during the year?"

"What the situation has been in the PC market?"

"How terrible the typhoons were last autumn?"



Concise, plain text

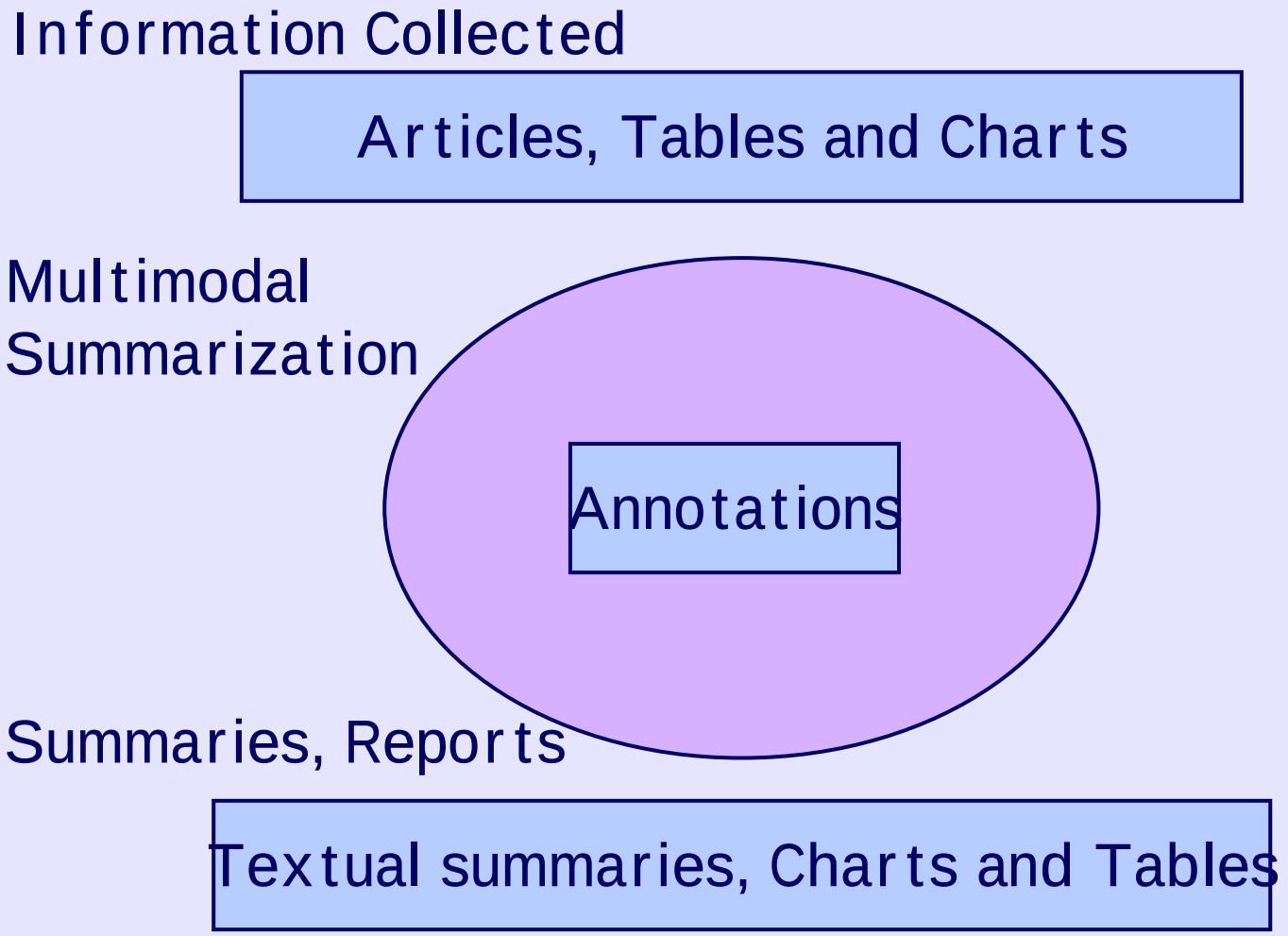
Information graphics

Multimedia presentation

text including references to graphics

graphics annotated with text

The Roles of Data Set



Course, Future Plan

- Announced in Nov. 2004
- Data set was started to deliver in Jan./April
- Round table meeting in Nov. 2005
- Workshop in March 2006

- 14 participants
- Two MLs:
  - closed to the participants
  - open widely for interested persons (about 90)

Contact Info & Online Proceedings

Documents used are Asian Languages but **participation from all over the world is more than welcome!!**

Open Submission Session for NTCIR-5

Inquiries: Noriko Kando at kando (at) nii. ac.jp

Online proceedings, application & other info: **http://research.nii.ac.jp./ntcir/**

Thanks Merci

Danke schön Gracie

Gracias Ta! Tack

Köszönöm Kiitos

Terima Kasih Khap Khun

Ahsante Tak

謝謝 ありがとう

http://research.nii.ac.jp/ntcir/