Overview of the 7th NTCIR Workshop

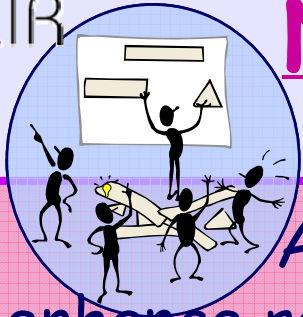**Noriko Kando**

National Institute of Informatics, Japan

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

kando (at) nii. ac. Jp

With thanks for **Tetsuya Sakai** for the slides

NTCIR: NII Test Collection for Information Retrieval

Research Infrastructure for Evaluating IA

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

■ Data sets, evaluation methodologies, and forum

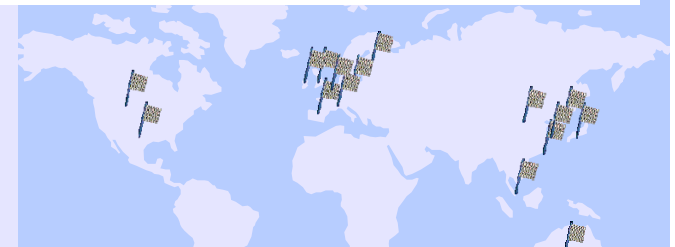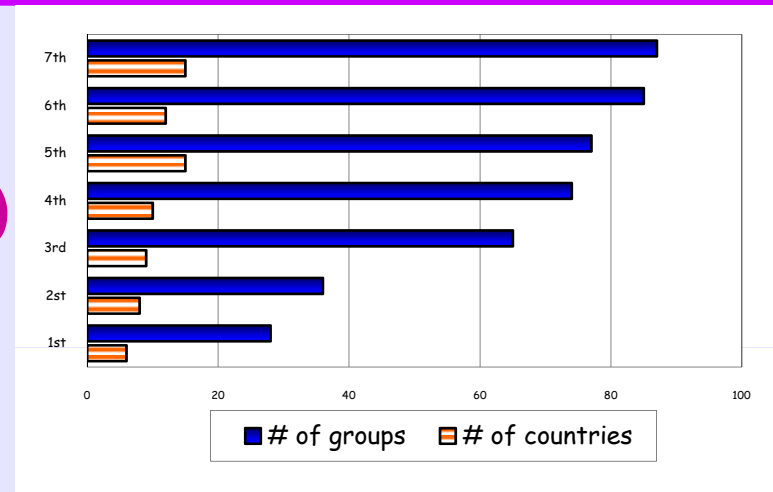## Project started in late 1997

- Once every 18 months

Data sets (Test collections or TCs)

- Scientific, news, **patents**, and web

- Chinese, Korean, Japanese, and English

Tasks

- IR: Cross-lingual tasks, patents, web,
- QA: Monolingual tasks, cross-lingual tasks
- Summarization, trend info., patent maps
- Opinion analysis, text mining

Community-based Research Activities

Noriko Kando

NTC7 OV 2008-12-17

Chart legend: ■ # of groups □ # of countries (categories: 1st, 2st, 3rd, 4th, 5th, 6th, 7th; axis 0 to 100)

NTCIR-7 participants
82 groups from 15 countries

Information access (IA)

- Whole process ofpreparing information from the vast collection of documents usable by users.

- For example, IR, text summarization, QA, text mining, and clustering

- Use human assessments as success criteria

Focus of NTCIR

Lab-type IR Test

Asian Languages/cross-language

Variety of Genre
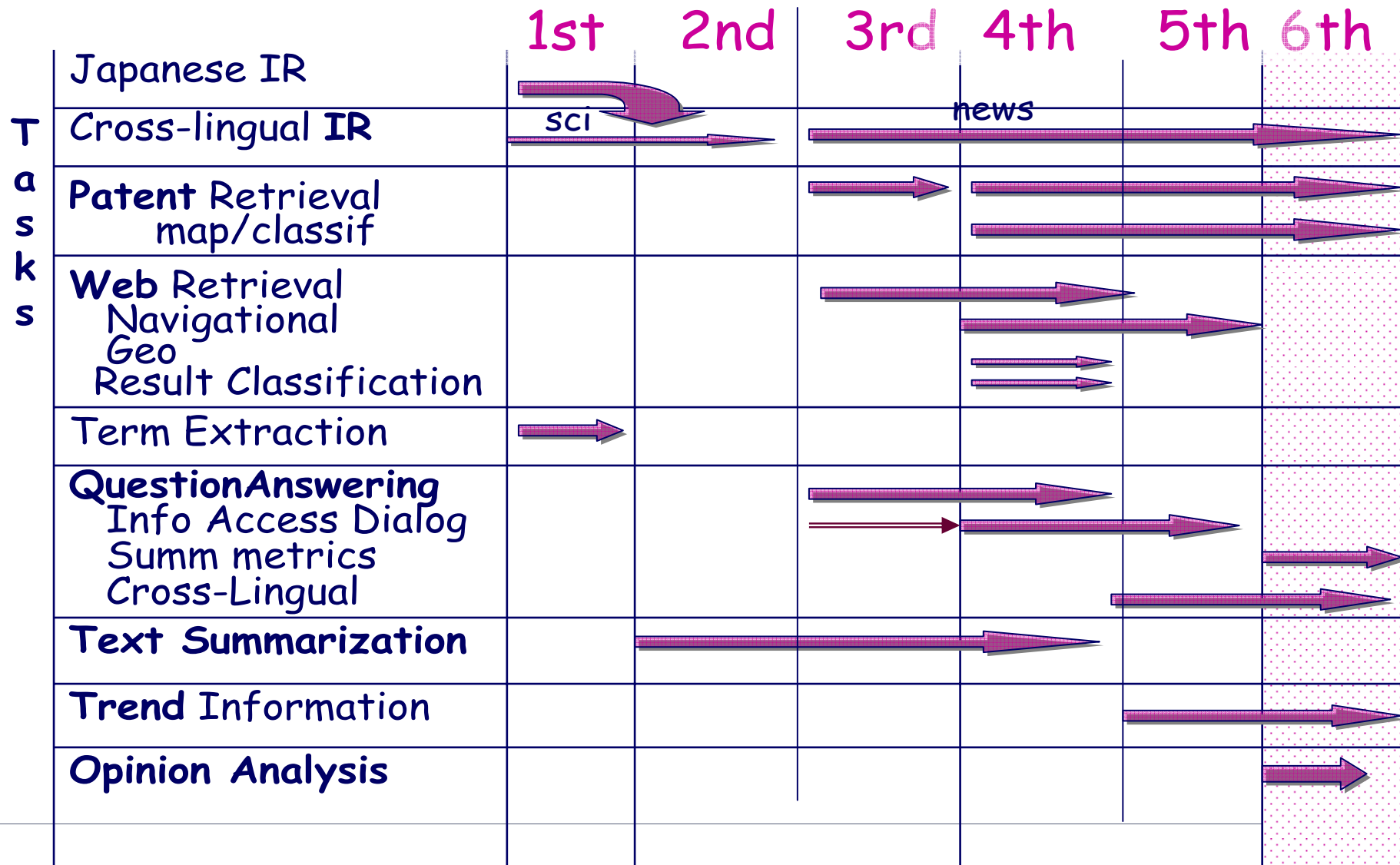
Parallel/comparable Corpus

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 at past NTCIRs

Tasks	1st	2nd	3rd	4th	5th	6th
Japanese IR						
Cross-lingual **IR**	sci		news			
Patent Retrieval						
map/classif						
Web Retrieval						
Navigational						
Geo						
Result Classification						
Term Extraction						
QuestionAnswering						
Info Access Dialog						
Summ metrics						
Cross-Lingual						
Text Summarization						
Trend Information						
Opinion Analysis						

NTCIR-7 Clusters

Cluster 1. Advanced CLIA

- Complex CLQA (Chinese, Japanese, English)
- IR for QA (Chinese, Japanese, English)

Cluster 2. User-Generated :

- Multilingual Opinion Analysis

Cluster 3. Focused Domain : Patent

- Patent Translation ; English -> Japanese,
- Patent Mining paper -> IPC

Cluster 4. MuST :

- Multi-modal Summarization of Trends

MuST; Visualization Challenge

NTCIR-7 is made up of...

- Cluster 1: Advanced Cross-lingual Information Access (ACLIA) = CCLQA + IR4QA
- Cluster 2: Multilingual Opinion Analysis task (MOAT) + CLIRB
- Cluster 3: Focused Domains
  = PATMT + PATMN
- Multimodal Summarization of Trend information (MuST)
- The 2nd International Workshop on Evaluating Information Access (EVIA)
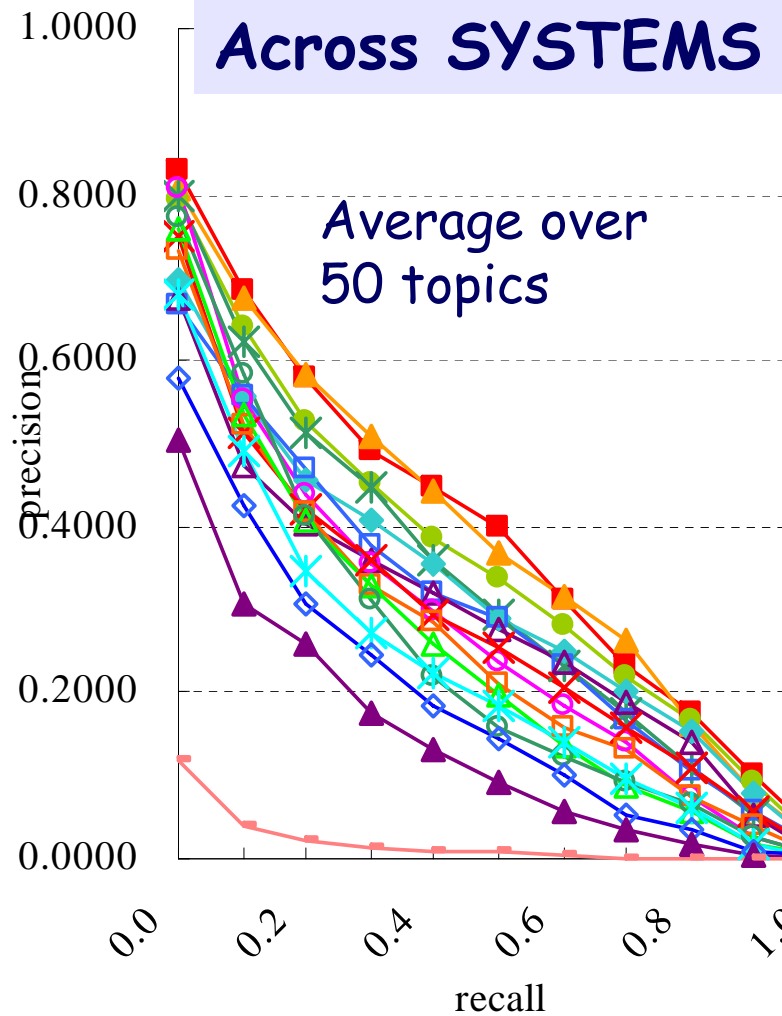
Evaluation Workshops

- "evaluation"

  It is not an competition!   not an exam!

- Constructs a common data set usable for experiments.

- provides to participants the data sets and unified procedures for evaluation

  - Each participating research group conducts experiments with various approaches and can participate with own purpose.

- Successful examples; TREC, CLEF, DUC, INEX, and TAC, FIRE (new!)   Community-based activities

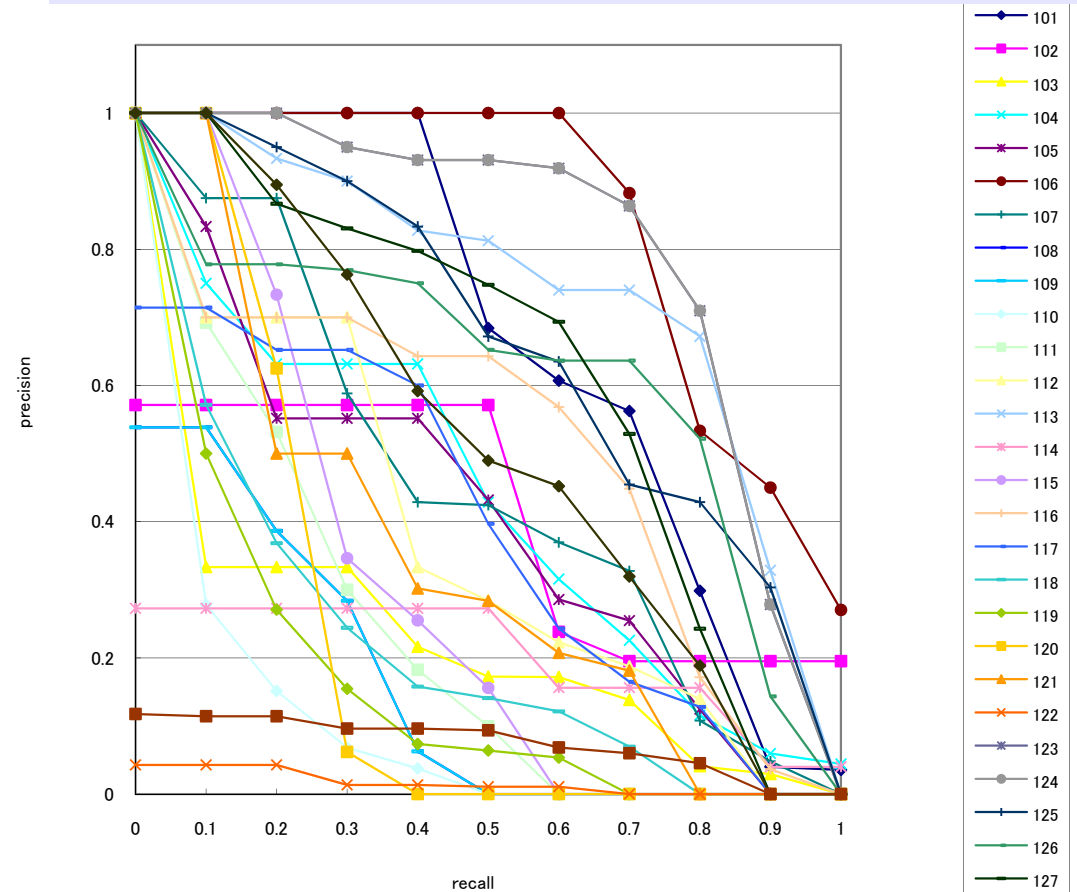- Implications are various

IA Systems Evaluation

- Engineering Level: Efficiency

- Input Level: ex. Exhaustivity, quality, novelty of DB

⇒ Process Level: Effectiveness ex. recall,precision

- Output Level: Display of output

- User Level: ex. Effort that users need

- Social Level: ex. Importance   (Cleverdon & Keen 1966)

Retrieval Difficulty Varies with Topics

Effectiveness Across SYSTEMS

Average over 50 topics

precision / recall

Effectiveness Across TOPICS on a System

precision / recall

Legend: 101, 102, 103, 104, 105, 106, 107, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123, 124, 125, 126, 127
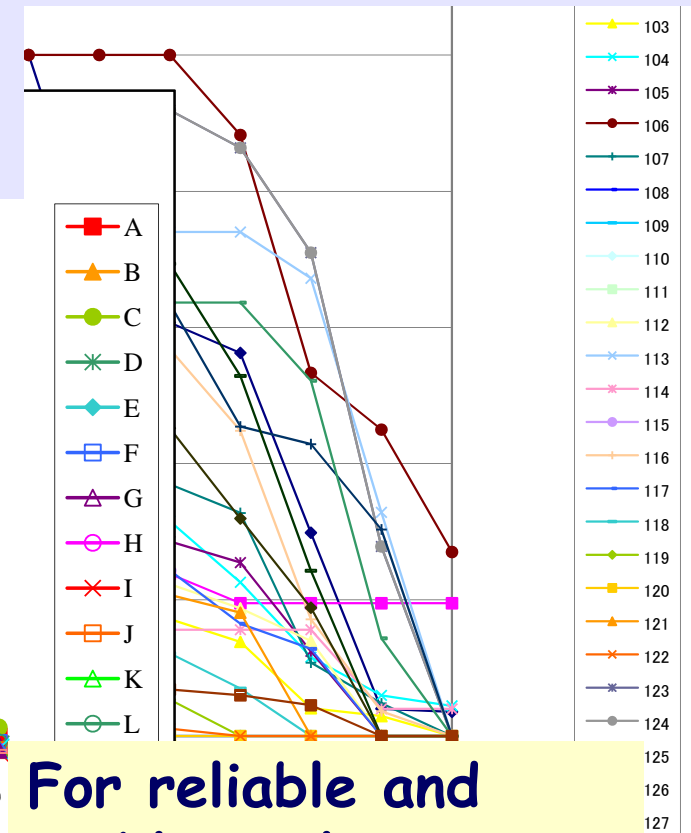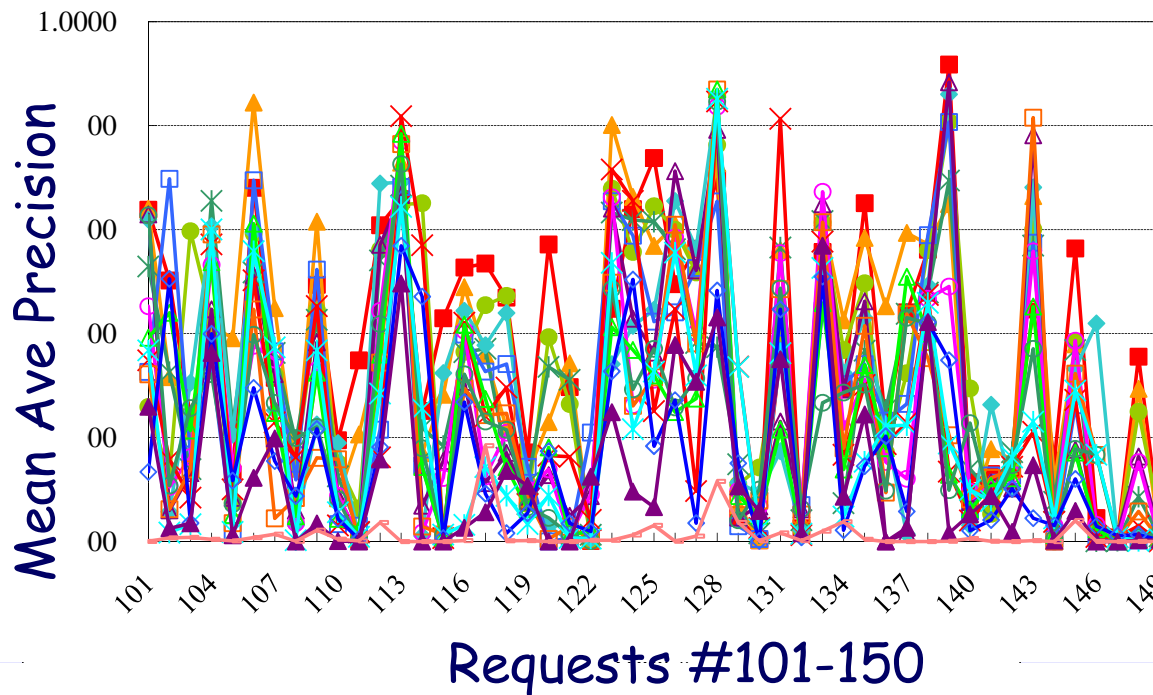
Retrieval Difficulty Varies with Topics

Effectiveness Across SYSTEMS

Effectiveness Across TOPICS on a System

"Difficult Topics" Vary with Systems

Legend (right chart):
103, 104, 105, 106, 107, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123, 124, 125, 126, 127

Legend (center chart): A, B, C, D, E, F, G, H, I, J, K, L

Y-axis: Mean Ave Precision

1.0000

X-axis: Requests #101-150

101 104 107 110 113 116 119 122 125 128 131 134 137 140 143 146 149

For reliable and stable evaluation, using substantial # topics is inevitable.

TC usable to evaluate?

Pharmaceutical R & D

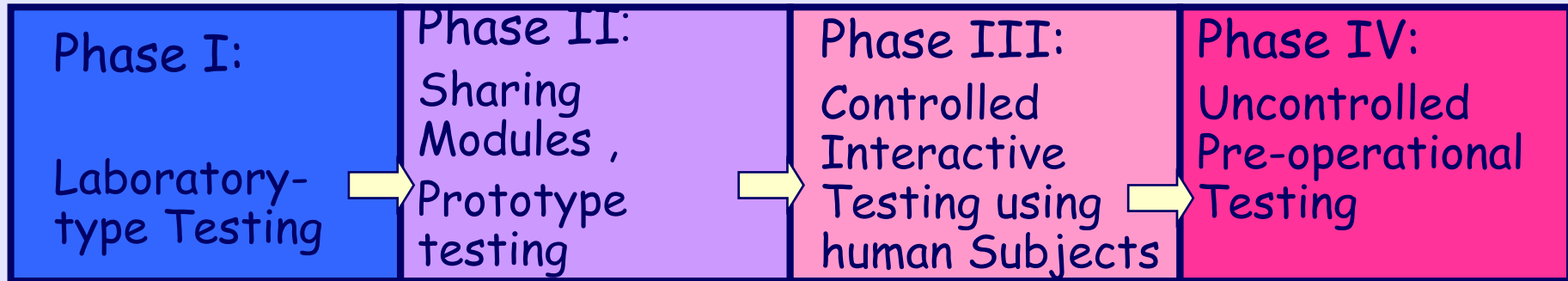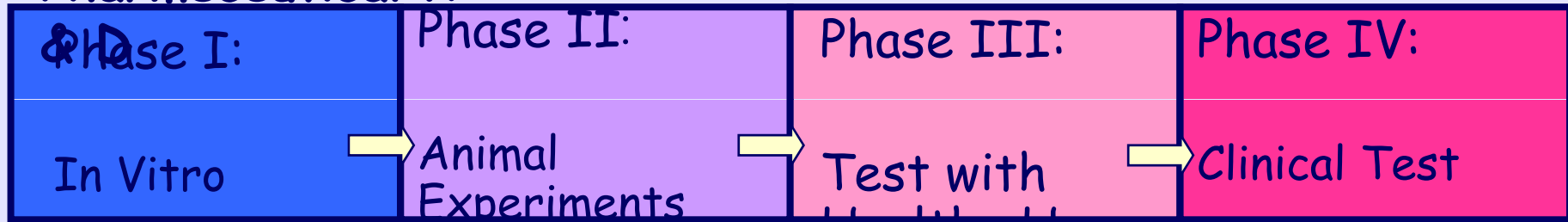| Phase I: In Vitro	Phase II: Animal Experiments	Phase III: Test with Healthy Human Subject	Phase IV: Clinical Test
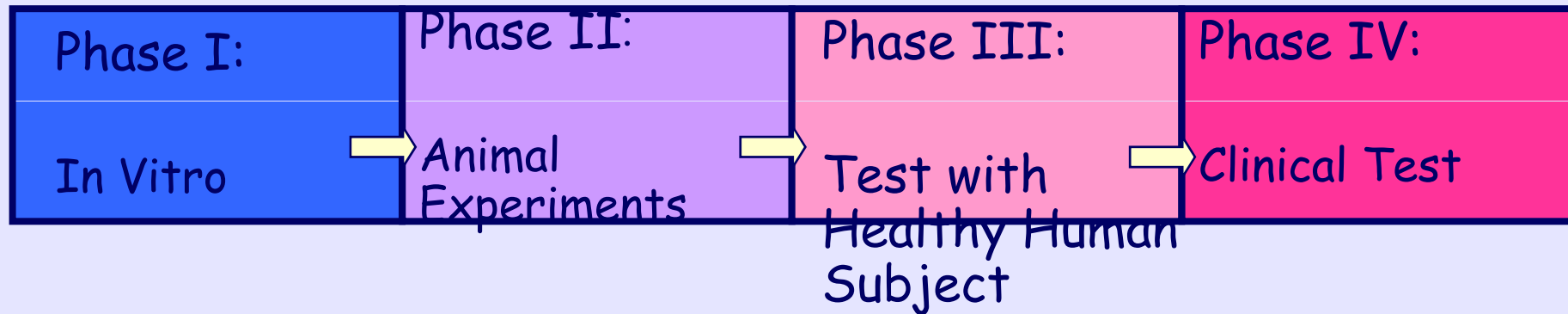
TC usable to evaluate what?

Test Collections

Users' information seeing tasks

Phase I: Laboratory-type Testing	Phase II: Sharing Modules, Prototype testing	Phase III: Controlled Interactive Testing using human Subjects	Phase IV: Uncontrolled Pre-operational Testing

Pharmeceutical R & D

Phase I: In Vitro	Phase II: Animal Experiments	Phase III: Test with Healthy Human Subject	Phase IV: Clinical Test

Levels of Evaluation

2.Input Level、6.Social Level

4.User Level、5.Output Levle

1.Engineering Level   efficiency

3.Process Level: effectiveness

Summary of "What is NTCIR"

- Providing a scientific basis for understanding the effectiveness of automated information access technologies

- Leveraging the R&D and technology transfer

- Reusable Test collection is a key component

- Evaluating search effectiveness is not easy. A small-scale or carelessly-designed TCs may skew the test results
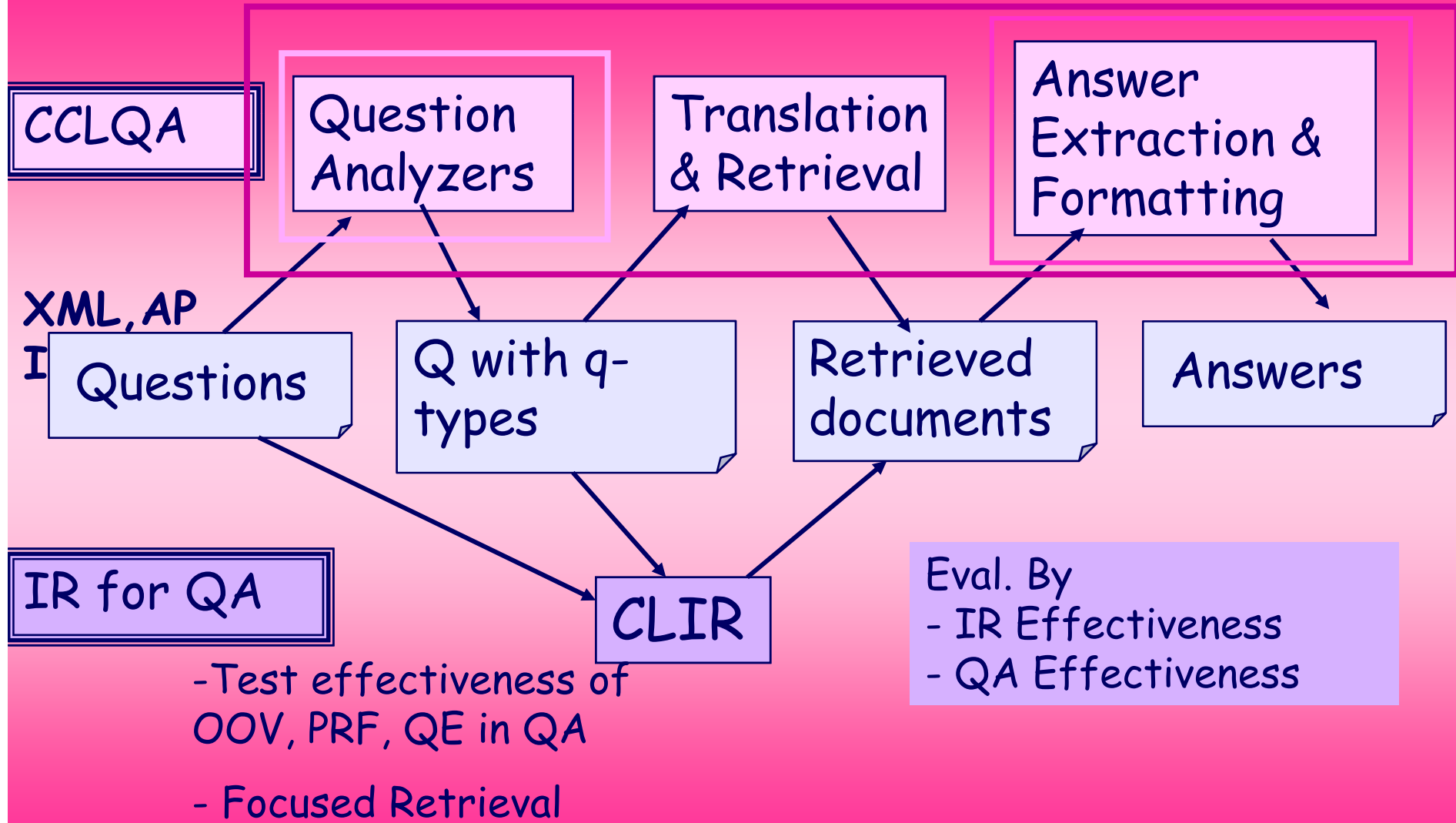
NTCIR-7: Advanced CLIA

Teruko Mitamura (CMU)
Eric Nyberg (CMU)

Ruihua Chen (MSRA)
Fred Gey (UCB),
Donghong Ji (Wuhan Univ)
Noriko Kando (NII)
Chin-Yew Lin (MSRA)
Chuan-Jie Lin (Nat Taiwan Ocean Univ)
Tsuneaki Kato (Tokyo Univ)
Tatsunori Mori (Yokohama N Univ)
Tetsuya Sakai (NewsWatch)

Advisor: K.L.Kwok (Queen College)

NTCIR-7: **Advanced CLIA**

CCLQA

Question Analyzers

Translation & Retrieval

Answer Extraction & Formatting

XML, API

Questions

Q with q-types

Retrieved documents

Answers

IR for QA

CLIR

Eval. By
- IR Effectiveness
- QA Effectiveness

-Test effectiveness of OOV, PRF, QE in QA

- Focused Retrieval

ACLIA
Complex Cross-lingual Question Answering (CCLQA) Task

Sample System

CLQA System

QA participant can decide whether to collaborate with CLIR systems, or develop an IR sub-system.

③ Question Analysis → Document Retrieval → Answer Extraction → Answer Selection

XML: Question | Analyzed Result | Retrieved Result | Extracted Result | Final Answer

CLIR System ①

② ③ Document Retrieval

② In collaboration with CLQA, a CLIR system can also use translated keyterms and answer type analysis.

③ Translation often happens in here.

IR Evaluation

QA Evaluation

Small teams that do not possess an entire QA system can contribute

Different teams can exchange and create a "dream-team" QA system

IR and QA communities can collaborat

CCLQA= Complex CLQA

- Moving towards Advanced Complex Questions from Factoid Questions (NTCIR-5, NTCIR-6)
- 4 questions types (events, biographies, definitions, and relationships)
- Examples of Complex Questions
  - **Definition questions**: *What is the Human Genome Project?*
  - **Relationship questions**: *What is the relationship between Saddam Hussein and Jacques Chirac?*
  - **Event questions**: *List major events in formation of European Union.*
  - **Biography questions**: *Who is Kim Jong-Il?*

CCLQA = Complex Cross-lingual QA

- Three document languages:

Simplified Chinese (CS)

Traditional Chinese (CT)

Japanese (JA)

- Four question languages:

CS, CT, JA plus English (EN)

- Complex questions for Cross-Lingual QA

EN-CS, EN-CT and EN-JA

- Monolingual QA with the same complex questions

CS-CS, CT-CT and JA-JA

- Combination system evaluation

QA teams using IR4QA runs from other teams

Evaluation Metrics

- "Manual" evaluation

Nugget pyramid method [Lin/Demner-Fushman 06] using multiple assessors for judging nugget matches (Weighted F-measure)

- Automatic evaluation

POURPRE [Lin/Demner-Fushman 05] modified for Chinese and Japanese

ACLIA: Evaluation EPAN tool
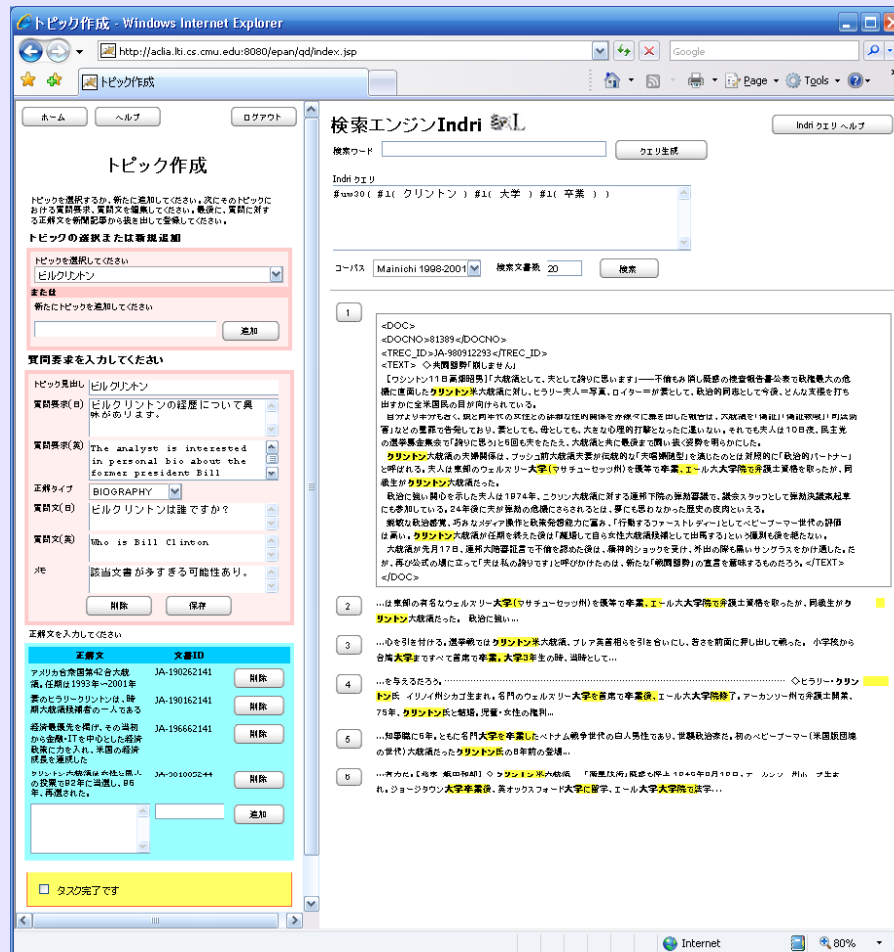
ACLIA: Evaluation EPAN tool
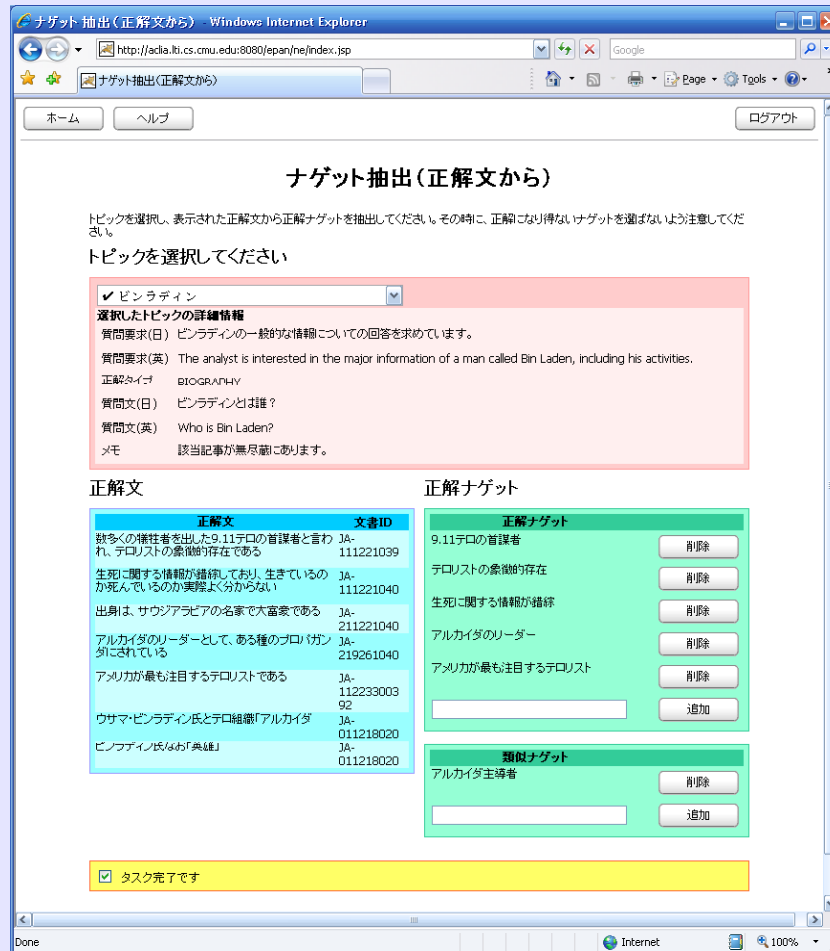
CCLQA:

Nugget Pyramid

IR4QA:

MAP

MS nDCG

Q-Measure

(preference-based)

Traditional "ad hoc" IR vs IR4QA

- Ad hoc IR (evaluated using *Average Precision* etc.)
- Find as many (partially or marginally) relevant documents as possible and put them near the top of the ranked list
- IR4QA (evaluating using... *WHAT?*)
- Find relevant documents containing different correct answers?
- Find multiple documents supporting the same correct answer to enhance reliability of that answer?
- Combine partially relevant documents A and B to deduce a correct answer?

Average Precision (AP)

$$AP = \frac{1}{R} \sum_r I(r) \frac{C(r)}{r}$$

Precision at rank r

Number of relevant docs

1 iff doc at r is relevant

- **Used widely since the advent of TREC**
- **Mean over topics is referred to as "MAP"**
- **Cannot handle graded relevance**
(but many IR researchers just love it)

Q-measure (Q)

$$Q\text{-}measure = \frac{1}{R} \sum_r I(r) \frac{C(r) + \beta cg(r)}{r + \beta cg^*(r)}$$

- **Generalises AP and handles graded relevance**
- **Properties similar to AP and higher discriminative power**
- **Not widely-used, but has been used for QA and INEX as well as IR**

Blended ratio at rank r (Combines Precision and normalised Cumulative Gain)

Sakai and Robertson EVIA 08 provides a user model for AP and Q

nDCG (Microsoft version)

Sum of discounted gains for a system output

$$nDCG = \frac{\sum_{r=1}^{l} g(r)/\log(r+1)}{\sum_{r=1}^{l} g^*(r)/\log(r+1)}$$

Sum of discounted gains for an *ideal* output

- **Fixes a bug of the original nDCG**
- **But lacks a parameter that reflects the user's persistence**
- **Most popular graded-relevance metric**

IR4QA evaluation package
(Works for ad hoc IR in general)

Computes
AP, Q, nDCG,
RBP, NCU
[Sakai and Robertson
EVIA 08] and so on
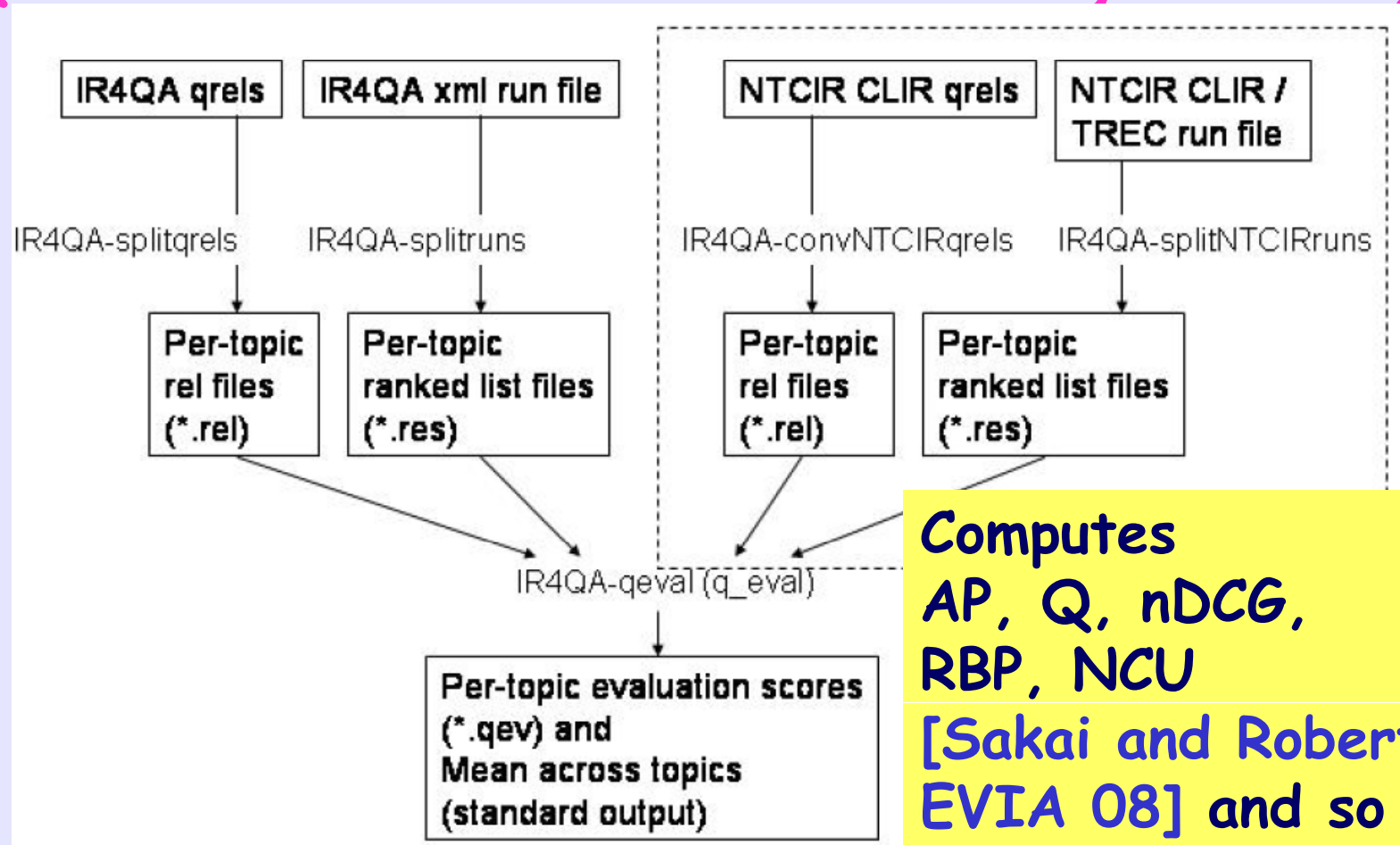
http://research.nii.ac.jp/ntcir/tools/ir4qa_eval-en

Table 1. IR4QA participants.

team name	organisation
BRKLY	University of California, Berkeley
CMUJAV	Language Technologies Institute, Carnegie Mellon University
CYUT	Chaoyang University of Technology
HIT	Heilongjiang Institute of Technology User Group: HIT2 NLP Joint Lab
KECIR	Shenyang Institute of Aeronautical Engineering
MITEL	Institute of Computing Technology, Chinese Academy of Sciences
NLPAI	College of Computer Science and Technology, Wuhan University of Science and Technology
NTUBROWS	CSIE, National Taiwan University
OT	Open Text Corporation
RALI	University of Montreal
TA	Toyohashi University of Technology
WHUCC	Computer Center of Wuhan University

- **12 participants from China/Taiwan, USA, Japan**
- **40 CS runs (22 CS-CS, 18 EN-CS)**
- **26 CT runs (19 CT-CT, 7 EN-CT)**
- **25 JA runs (14 JA-JA, 11 EN-JA)**

Monolingual   Crosslingual

Oral presentations

- CMUJAV (CS-CS, EN-CS, JA-JA, EN-JA)
- Proposes Pseudo Relevance Feedback using Lexico-Semantic Patterns (LSP-PRF)
- CYUT (EN-CS, EN-CT, EN-JA)
- Uses Wikipedia in several ways; post hoc results
- MITEL (EN-CS, CT-CT)
- SMT and Baidu used for translation; data fusion
- RALI (CS-CS, EN-CS, CT-CT, EN-CT)
- Uses Wikipedia in several ways; high performance after bug fix

Other interesting approaches

- BRKLY (JA-JA) A very experienced TREC/NTCIR participant
- HIT (EN-CS) PRF most successful
- KECIR (CS-CS) **Query expansion length optimised for each question type** (definition, biography…)
- NLPAI (CS-CS) **Uses question analyses files from other teams   (next slide)**
- NTUBROWS (CT-CT) Query term filtering, data fusion
- OT (CS-CS, CT-CT, JA-JA) Data fusion-like PRF
- TA (EN-JA) SMT document translation from NTCIR-6
- WHUCC (CS-CS) Document reranking

Please visit the posters of all 12 IR4QA teams!

NLPAI (CS-CS) used question analysis files from other teams.

```
CSWHU-CS-CS-01-T:
<KEYTERMS>
<KEYTERM SCORE="1.0">宇宙大爆炸</KEYTERM>
<KEYTERM SCORE="0.3">理论</KEYTERM>
</KEYTERMS>
Apath-CS-CS-01-T:
<KEYTERMS>
<KEYTERM SCORE="1.0">宇宙大爆炸理论</KEYTERM>
</KEYTERMS>
CMUJAV-CS-CS-01-T:
<KEYTERMS>
<KEYTERM SCORE="1.0">宇宙</KEYTERM>
<KEYTERM SCORE="1.0">大</KEYTERM>
<KEYTERM SCORE="1.0">爆炸</KEYTERM>
<KEYTERM SCORE="1.0">理论</KEYTERM>
<KEYTERM SCORE="1.0">宇宙 大 爆炸 理论</KEYTERM>
<KEYTERM SCORE="1.0">宇宙大爆炸理论</KEYTERM>
<KEYTERM SCORE="1.0">宇宙 大 爆炸</KEYTERM>
<KEYTERM SCORE="1.0">宇宙大爆炸</KEYTERM>
</KEYTERMS>
```
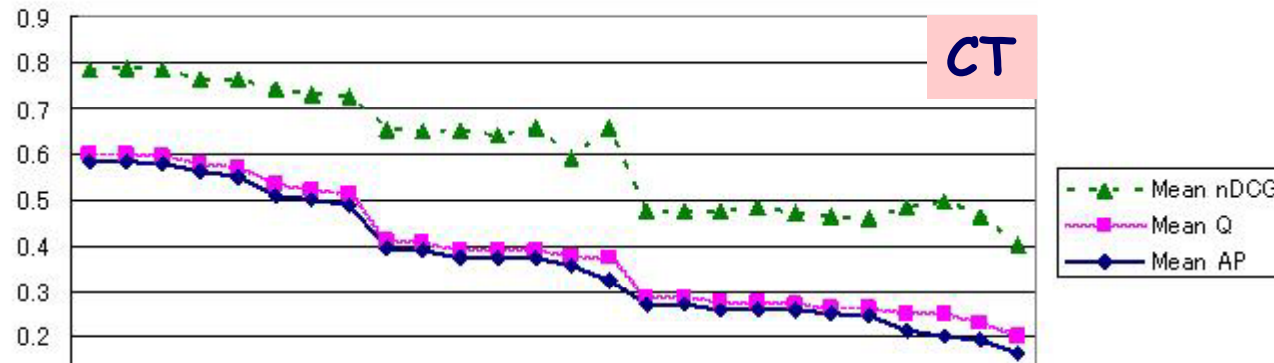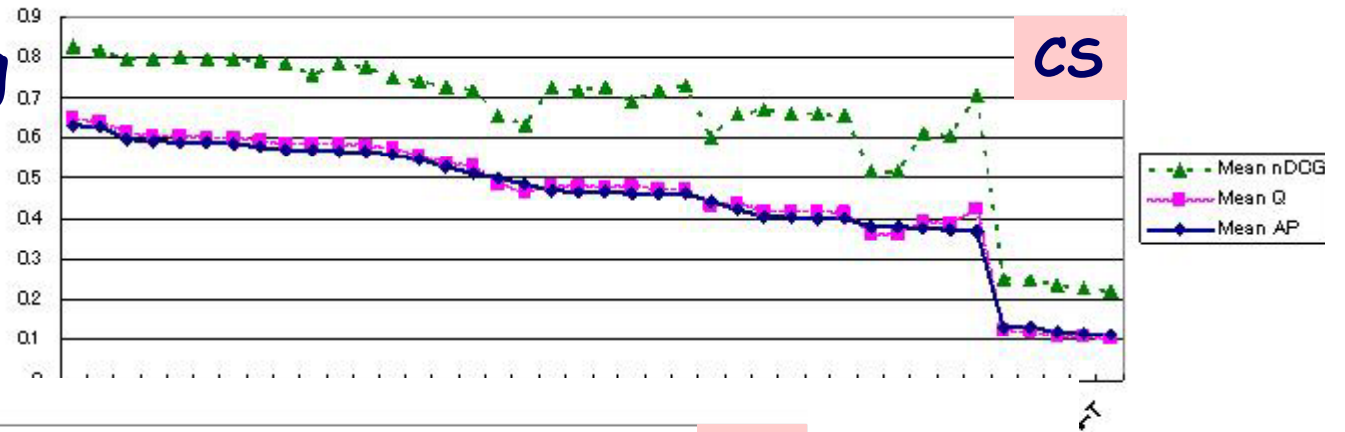
Different teams come up with different set of query terms with different weights. This clearly affects retrieval performance.

Special thanks to Maofu Liu (NLPAI)

System ranking by Q/nDCG vs that by AP
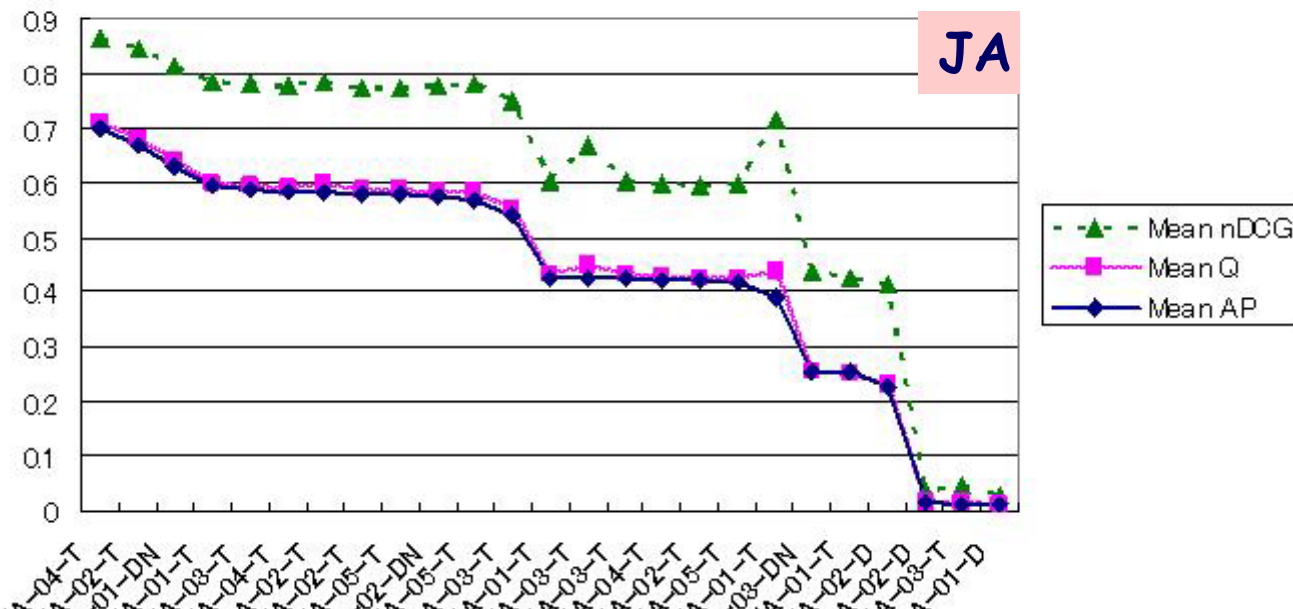
By definition, nDCG is more forgiving for low-recall runs than AP and Q.

Forming pseudo-qrels

QUESTION: Can we get away with not doing any relevance assessments at all?

1. Sort pooled docs by

(1) Number of runs that retrieved it; and then

(2) Sum of its ranks within these runs.

2. Take the top 10 docs in the sorted pool and treat them all as L1-relevant!

Very interesting results will be presented at NTCIR-7!

NTCIR-7: UGC (Blog)

David K Evans (NII -> Amazon Japan)
Yohei Seki (Toyohashi U Tech -> Columbia U)

LunWei Ku (National Taiwan Univ)
Le Sun (Chinese Academy of Science)
Hsin-Hsi Chen (National Taiwan Univ)
Noriko Kando (NII)

Opinion Analysis - Roadmap

Genre	Subjectivity	Holder	Polarity	Strength
News	NTCIR-6	NTCIR-6	NTCIR-6	
Review	NTCIR-7	NTCIR-7	NTCIR-7	NTCIR-7
Blog	NTCIR-8	NTCIR-8	NTCIR-8	NTCIR-8

Stakeholder	Temporal	Language	Granuality	Application
		Chinese	single-sent	Summarization
NTCIR-7		English	clause	QA
NTCIR-8	NTCIR-8	Japanese	multi-sent	Opinion tracking
		CJE	document	Consistency checkir
				Trend

Chinese, Japanese, English

NTCIR-7: MOAT (on News)

- **Documents:**
  - NEWS   CCEJ
- **CLIR on Blog (CLIRB) Cancelled**
- **Multilingual Opinion Analysis (MOAT)**
  - TraditionalC,Simplifed C, J,E
  - selecting relevant documents from ~25 topics used in ACLIA
  - Following Roadmap, but change the genre
  - Relevant, Opinionated, Polarity (Pos, Neg, Nue), Holder, Stakeholder (Object), ??Strength??

MOAT Participants

Beijing university of posts and telecomunications
Chinese Academy of Sciences(NLPR-IACAS)
City University of Hong Kong
CUHK（The Chinese University of Hong Kong)-PolyU(The Hong Kong Polythechnic University)-Tsinghua(Tsinghua University)
DAEDALUS, S.A.
Dalian University of Technology
Hiroshima City University
Information and Communications University
Keio University
Louisiana State University(University of Maryland College Park)

National Taiwan University
NEC
NEU Natural Language Processing Lab
Peking University
Peking University(ICL)
Pohang University of Science and Technology
SICS - Swedish Institute of Computer Science
Technical University of Darmstadt
The Graduate University for Advanced Studies(SOKENDAI).
Tornado Technologies Co., Ltd., Taiwan.
Toyohashi University of Technology
University of Neuchatel
University of Sussex
Yuan Ze Univ.

80+ registerd, 30+ resigned when docs were changed, 42 registered to News MOAT, 24 sugmitted

NTCIR-7: Focused Domain (Patent)

Atsuhi Fujii (Univ Tsukuba)

Taiich Hashimoto (Tokyo Insti Tech)
Makoto Iwayama (Tokyo Insti Tech/ Hitach)
Hidetsugu Nanba (Hiroshima City Univ)
Masao Utiyama (NICT),
Mikio Yamamoto, U Tsukuba)
Takehito Utsuro (U Tsukuba)

NTCIR-7: Focused Domain (Patent)

Documents:

 10 Yrs Japanese Patent Application (NTCIR4-5)

 10 Yrs USTPO Patents (NTCIR6)

 Parallel Sentence Data (1.8 M sentences JE Pairs)

 Scientific Paper Abstracts (NTCIR 1-2)

Patent Translation (PATMT) MT is key for CLIR

 Training: 1993-2000, Test: 2001-2002   One Ref Trans good??

 Intrinsic Eval. ;BLEU, human assessments

 Extrinsic Eval: CLIR task-based

Patent Mining (PATMN) Cross-Genre PAT & Scientific

 Classify Paper Abstracts in to IPC Classes

 ML approach: Classsify Absts to IPC Class

 IR Apprach: use invalidity search system to find
relevant Patent, then assign IPCs to Paper Absts.

Patent classification and mining at NTCIR

Organizers:
Makoto Iwayama (Hitachi Ltd/Tokyo Institute of Technology)
Hidetsugu Nanba (Hiroshima City University)
Taiichi Hashimoto (Tokyo Institute of Technology)
Atsushi Fujii (University of Tsukuba)
Noriko Kando (National Institute of Informatics)

NTC7 OV 2008-12-17

Noriko Kando

Goal: Automatic generation of patent maps.

Example: Blue light-emitting diodes

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

Systems automatically identify rows and columns

History

- NTCIR-4 (2003-2004): Patent-map-creation subtask
  - Direct approach to creation of patent maps
  - Hard tasks and insufficient evaluation
- NTCIR-5 (2004-2005): Classification subtask
  - Categorize patents to pre-defined categories called F-terms (multi-faceted and structured)
  - Relatively small number of test documents
  - Evaluate only strict matches in F-term hierarchy
- NTCIR-6 (2006-2007): Classification subtask
  - Increased the number of documents and topics (108 topics)
  - Evaluate partial matches in F-term hierarchy
- NTCIR-7 (2007-2008): Mining subtask

Feasibility Study: automatic patent map generation at NTCIR-4 (2003-2004)

search topic

application

JAPIO abst

PAJ

retrieval

documents

topics and documents in NTCIR-3 collection

Patent-map creation = Multi-faceted patent clustering

visualization

classification

multi-dimensional matrix
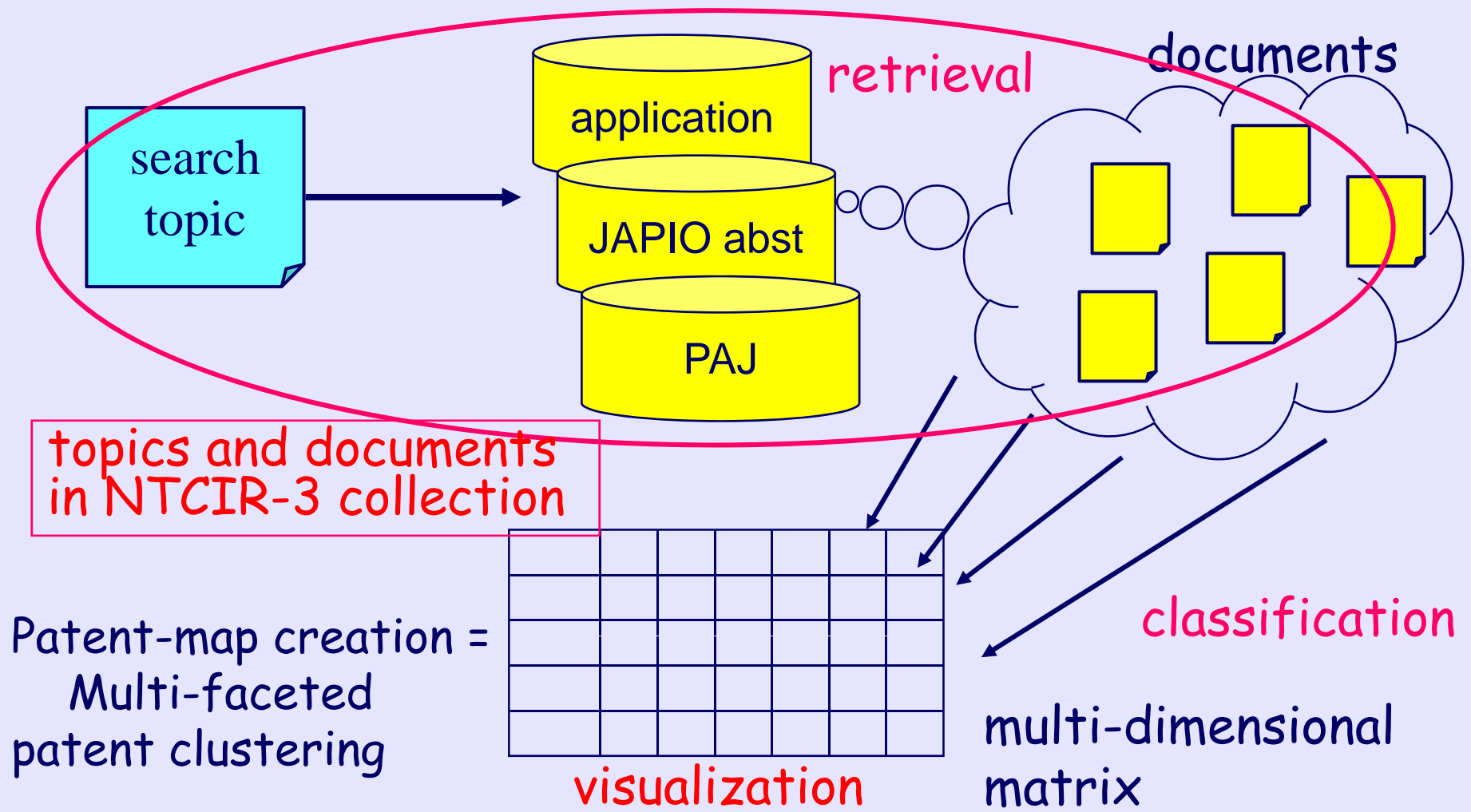
Classification task overview

Training data

Theme is given

Test data

Patents with themes and F-terms (1993-1997)

Training

5B001

Patents with themes and F-terms (1998-1999)

Sampling

PMGS (F-term descriptions)

Classifier

F-term assignment

5B001 AC04

Evaluation
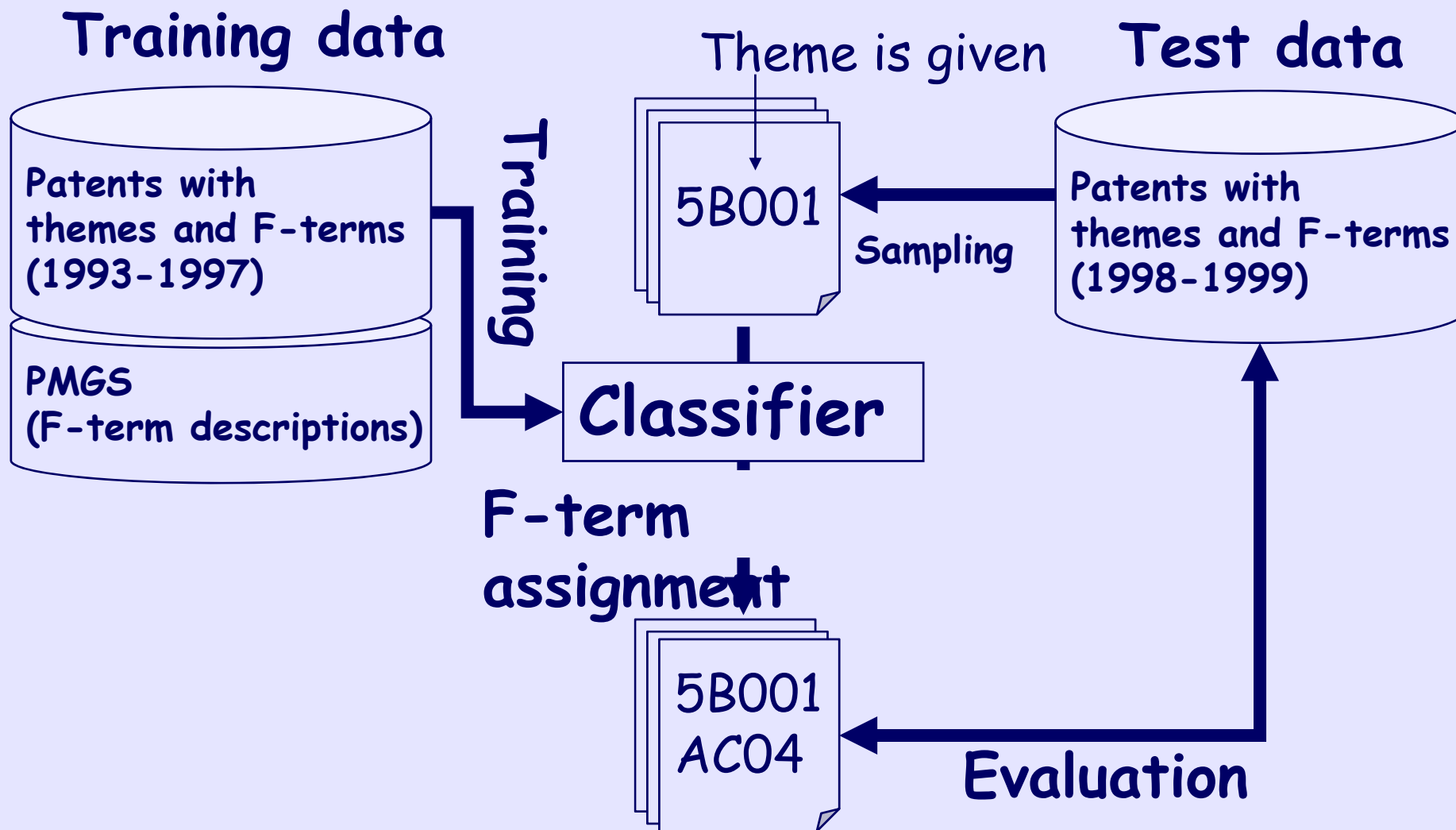
Patent mining at NTCIR-7 (2007-2008)

Searches and/or classifying patents and scientific papers into IPC

Research paper written in Japanese (Japanese / J2E subtasks)

Research paper written in English (English / E2J subtasks)

Japanese, English, and Cross-lingual (J-to-E, E-to-J) subtasks

A Participant System

Machine-translation module (E2J / J2E)

Text classification module

Patent data
- written in Japanese (Japanese / J2E)
- written in English (English / E2J)

List of IPC codes

Nanba, Fujii, Iwayama, and Hashimoto. "The Patent Mining Task in the Seventh NTCIR Workshop", Patent Information Retrieval Workshop at CIKM 2008 (2008)

Summary of patent classification and mining

- Automatic clustering of patents into "problems" and "solutions" are quite feasible, but labeling and controlled evaluation need more investigation.

- Granularity of F-term is appropriate for patent map creation and becoming good.

- Patent minting of scientific papers and patents are practically needed. n-KNN and machine learning have promise

  – The test collections for classification are available for research purpose. The one for mining will be available to the public after Workshop Meeting

Patent machine translation at NTCIR

Organizers:
Atsushi Fujii (University of Tsukuba)
Masao Utiyama (NICT)
Mikio Yamamoto (University of Tsukuba)
Takehito Utsuro (University of Tsukuba)

Fujii, Utiyama, Yamamoto, and Utsuro. "Toward the Evaluation of Machine Translation Using Patent Information", AMTA 2008

History of Patent IR at NTCIR

- NTCIR-3 (2001-2002)
  - Technology survey
    - Applied conventional IR problems to patent data

- NTCIR-4 (2003-2004)
  - Invalidity search
  - Addressed patent-specific IR problems

- NTCIR-5 (2004-2005)
  - Enlarged invalidity search

- NTCIR-6 (2006-2007)
  - Added English patents

2 years of JPO patent applications

* JPO = Japanese Patent O

5 years of JPO

Both document sets were published in 1993-2002

10 years of JPO patent applications

10 years of USPTO patents granted

* USPTO = US Patent & Trademark Offi

Patent machine translations at NTCIR-7 (2007-2008)

- Patent Machine Translation (MT) is realistic
  - Parallel corpora can potentially be produced from JPO/USPTO patent-document sets
  - Decoders for statistical MT (SMT) are available
- Two types of players
  - Organizer = Authors of this paper
    - Providing data, and evaluating participating MT systems
  - Participants = Research groups
    - They can use e.g., SMT and rule-based MT.
- Utility of patent MT
  - Cross-lingual patent retrieval
  - Filing patent applications in foreign countries

Producing parallel corpora

NTCIR

JPO applications 1993-2002 (3.5-M docs)

Comparable (not parallel)

USPTO grants 1993-2002 (1.3-M docs)

J E   J E   J E
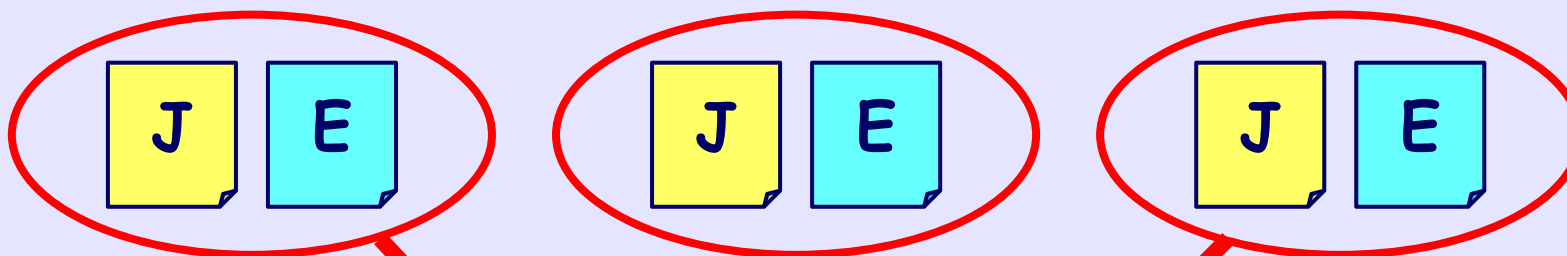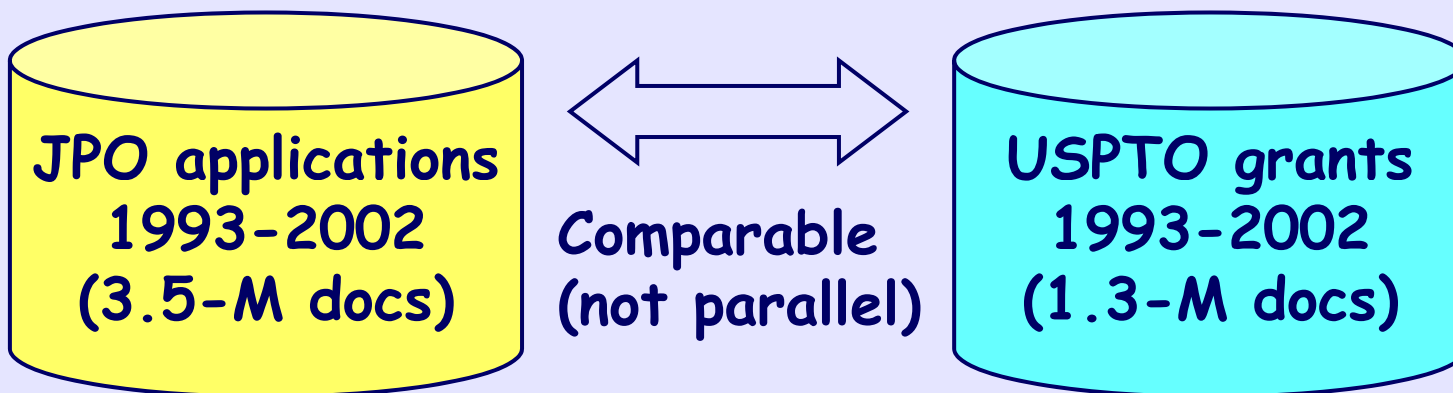
Sentence-alignment method [Utiyama and Isahara, 2007]

Targeting "background" and "description"

Sentence pairs

Patent family Patent set for same invention
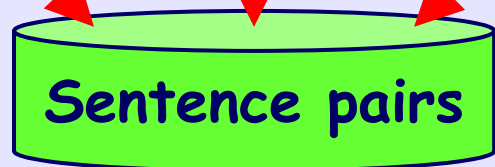
Parallel (alignment accuracy= 90%)

Extrinsic evaluation

NTCIR-5

Patent claim

Performed by organizers

Search topic in Japanese — *Human* → **Search topic in English**

Evaluation by BLEU

Training data 1.8-M sentence pairs → **MT system**

JPO applications 1993-2002

Invalidate patent

Translation in Japanese → **IR system**
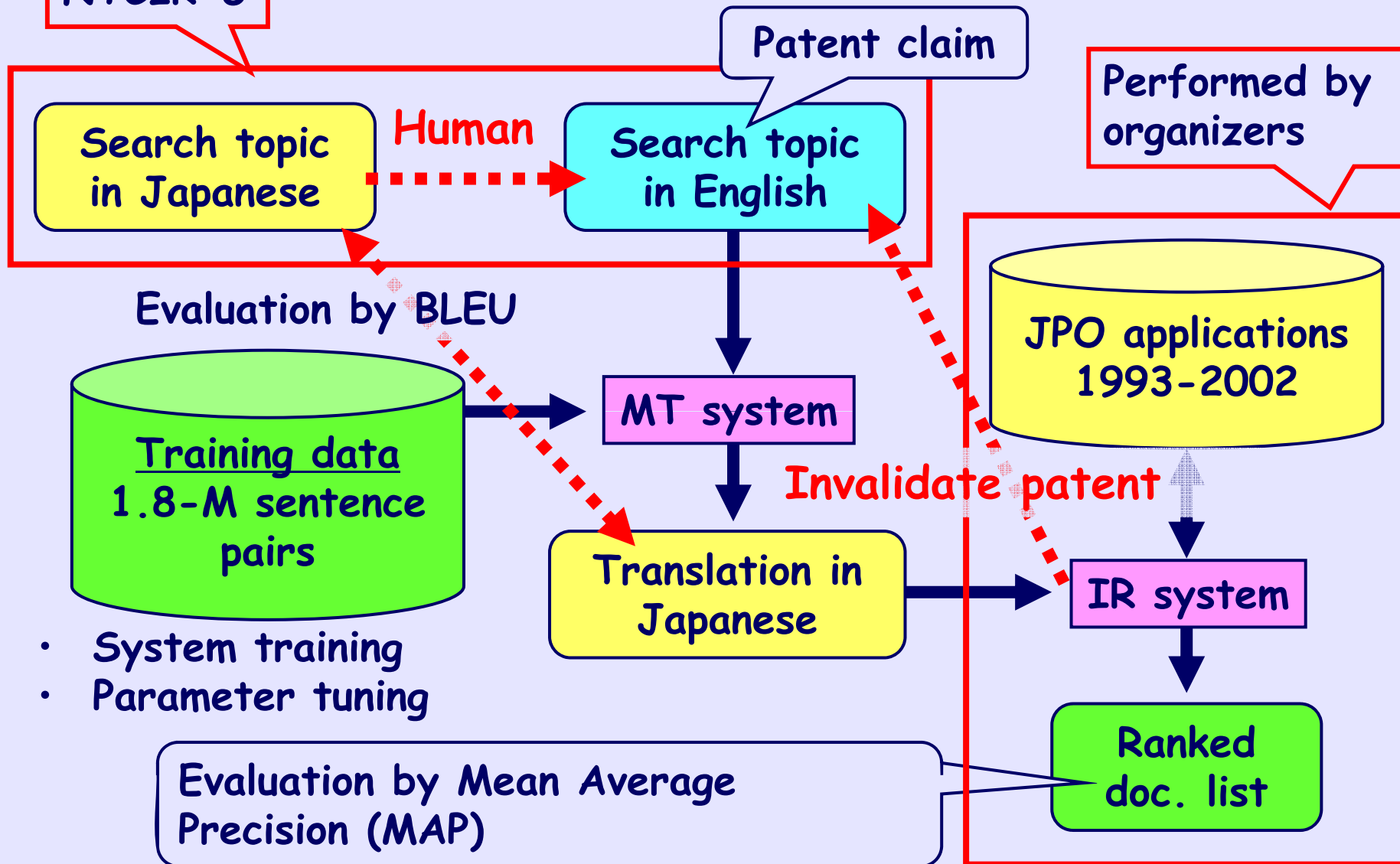
- System training
- Parameter tuning

Ranked doc. list

Evaluation by Mean Average Precision (MAP)

Patent families

- Member patents often claim "priority" under the Paris Convention
  - Related patents can easily be identified by priority numbers
  - 85 K patent families (J-to-E) were identified
- Merit of priority-based patent families
  - The application date is retroactive to the original date
  - First-to-file system in many countries

J | translating → | E

JPO
Aug 22, 2008

USPTO
Oct 24, 2008

claiming priority → Aug 22, 2008
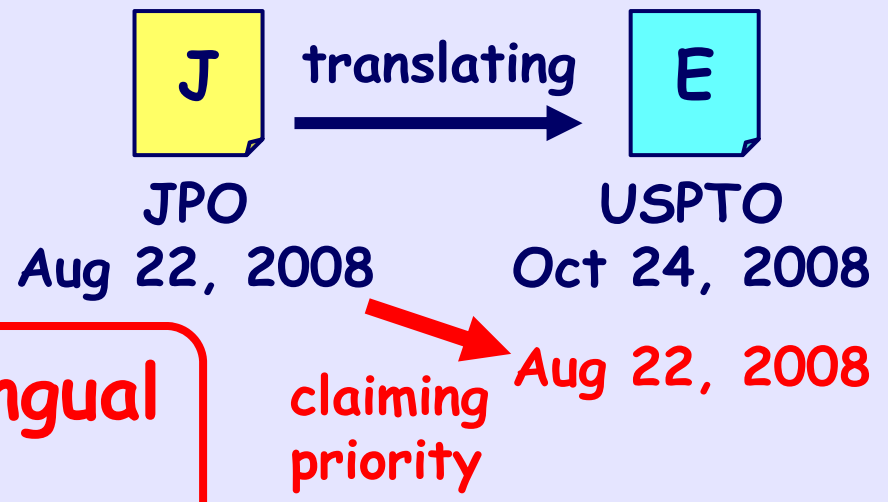
Free (or inexpensive) bilingual corpora are growing! ! !

Example of patent family

Invention related to Microactuators

JPO

USPTO

Patent family can be identified by priority number

J-E sentence pairs can be extracted from corresponding field

Evaluation Methods

- Intrinsic evaluation
    - Automatic evaluation by BLEU
    - Manual evaluation
        - Adequacy and Fluency by 5-point rating

- Extrinsic evaluation
    - Query translation for Cross-Lingual Patent Retrieval (CLPR), measured by Average Precision (AP)

Patent machine translation

- Constructed a large test collection for J/E MT: USTPO and JPO with 10 years of full texts
- Large-scale sentence-alignment dataset (E-J sentence pairs)
- Statistical MT (SMT)* vs. rule-based MT
- Results demonstrated:
  - SMT is much better for CLIR
  - Rule-based MT is good for human evaluations
  - Human evaluations and creation of reference translations must be carefully done (in the real world, professional patent translators do use MT).
- Test collection will be available for research purpose after the workshop meeting

*SMT : a system automatically learns the translation rules from the given large-scale sentence pairs.

MuST: Multimodal Summarization for Trend Information

Tsuneaki Kato (Tokyo Univ)
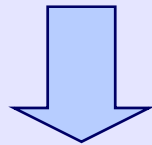## Mitsunori Matsushita (NTT Comm Sci Lab → Kansei Univ)

MuST

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?"
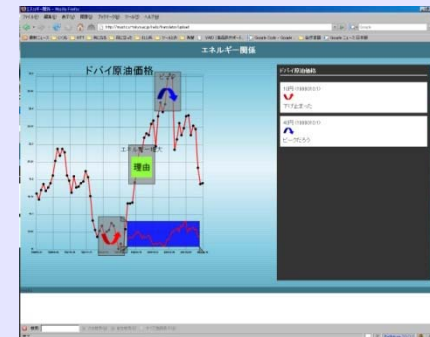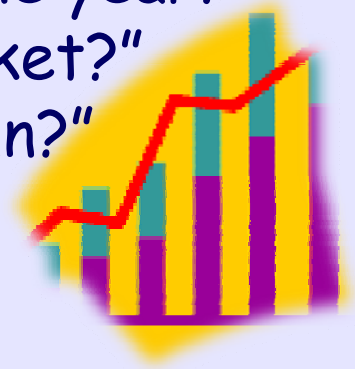"How terrible the typhoons were last autumn?"

Concise, plain text
Information graphics
Multimedia presentation
  text including references to graphics
  graphics annotated with text

Visualization Platform

The Roles of Data Set

Information Collected

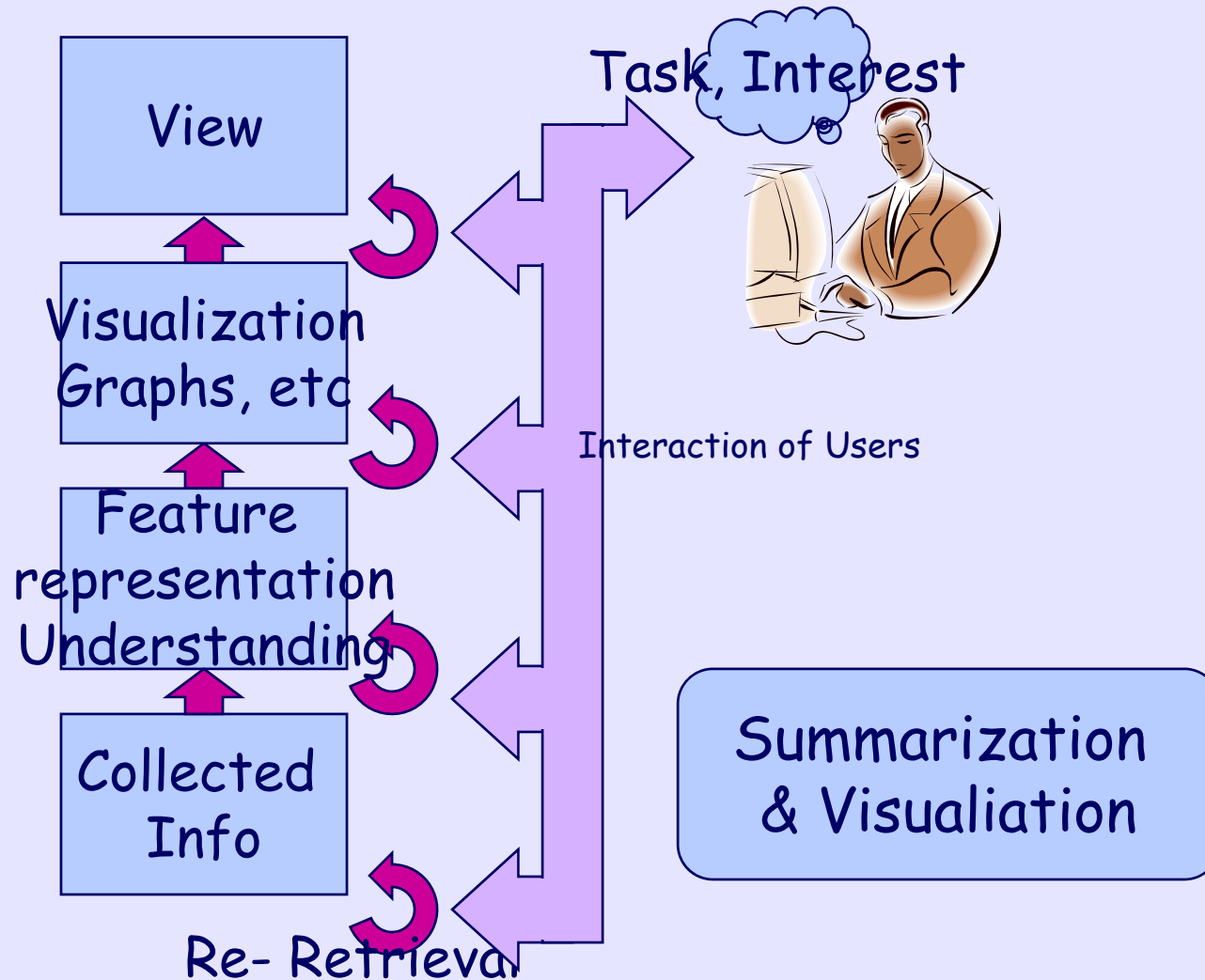Articles, Tables and Charts

Multimodal
Summarization

Visualization
software

Annotations

Summaries, Reports

Textual summaries, Charts and Tables

# Interactive and Exploratory Support of Information Utilization

View

Visualization Graphs, etc

Feature representation Understanding

Collected Info

Re- Retrieval

Task, Interest

Interaction of Users

Summarization & Visualiation

Noriko kando

Multimodal Summarization for Trend information (MuST)

Example: Visualising the Japanese cabinet support rate
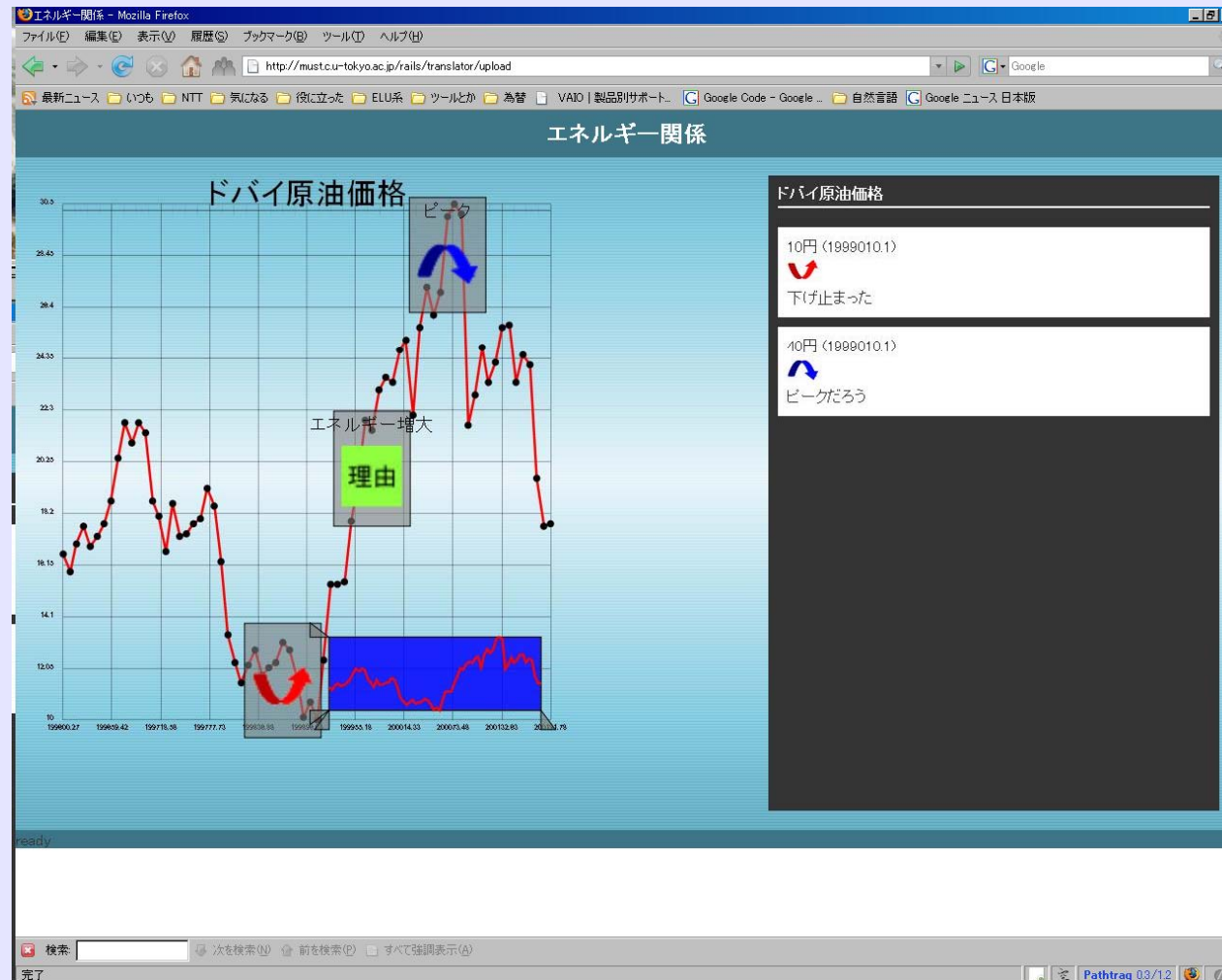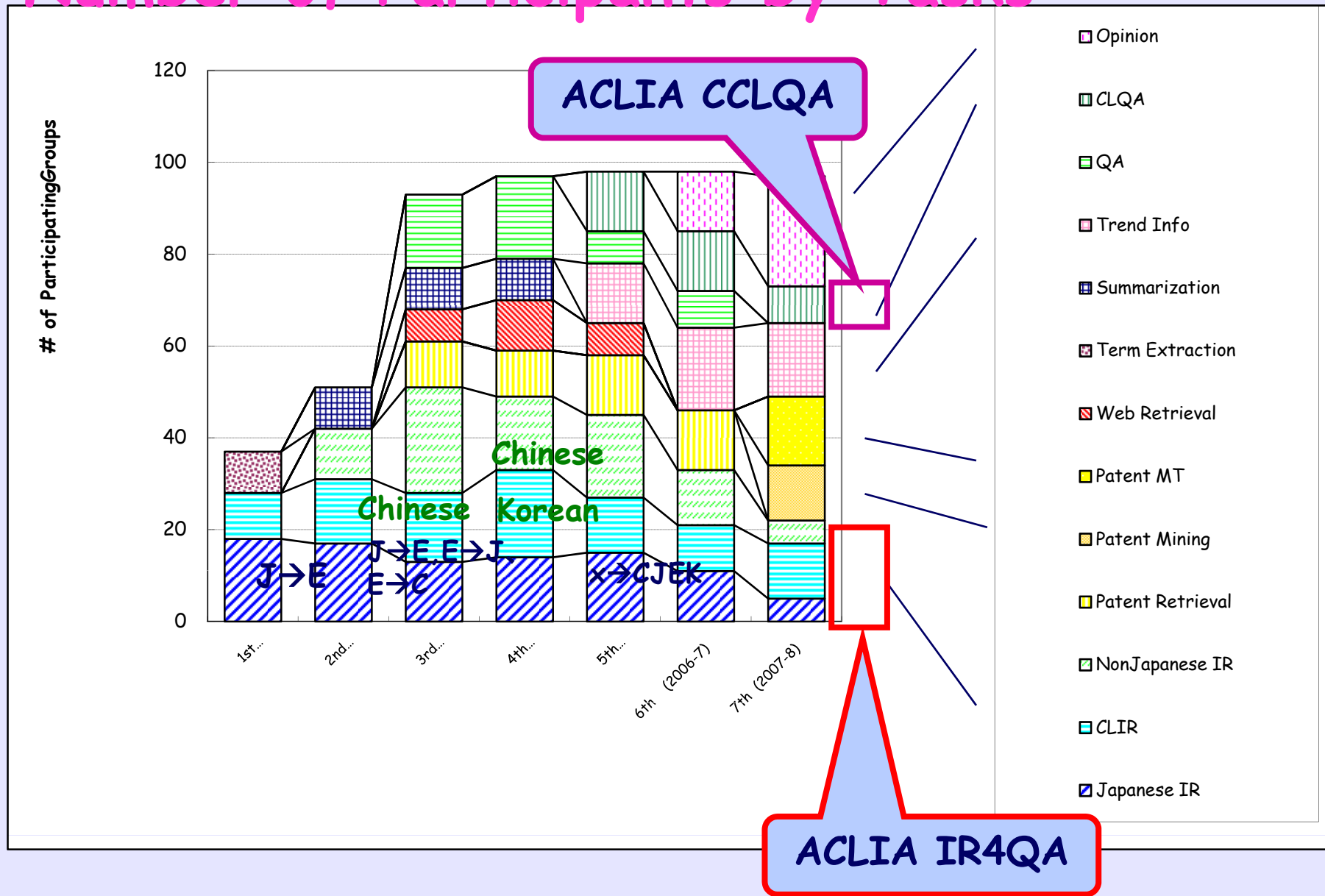
Gold standard
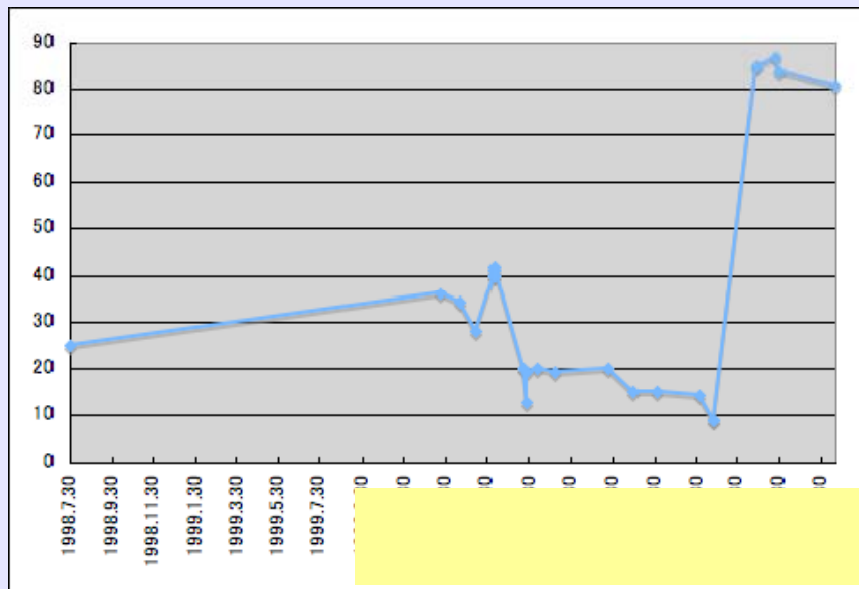
System output
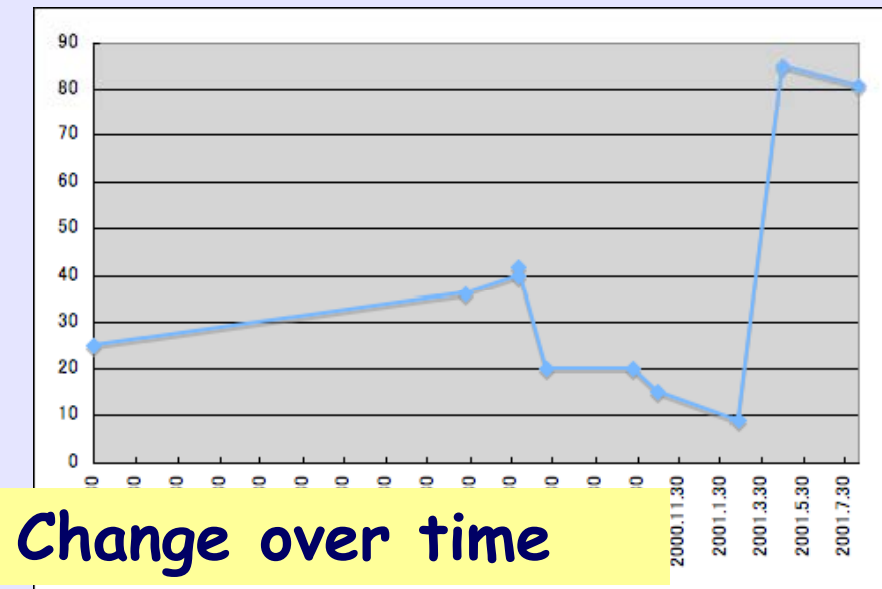
Change over time

Visualization Platform

Number of Participants by Tasks

[CCLQA]
- Academia Sinica
- Beijing Univ of Posts &   Telecoms, China
- Carnegie Mellon Univ
- NICT
- NTT Corporation
- Shenyang Institute of Aeronautical Engineering
- Wuhan Univ
- Yokohama National Univ

[IR4QA]
- Carnegie Mellon Univ
- Chaoyang Univ of Technology
- Chinese Academy of Sciences(ICT)
- Harbin Institute of Technology + Heilongjiang Institute of Technology
- National Taiwan Univ
- Open Text Corporation
- Shenyang Institute of Aeronautical Engineering
- Toyohashi Unive of Technology
- Univ of California, Berkeley
- Univ of Montreal
- Wuhan Univ
- Wuhan Univ of Science and Technology

[MOAT]
- Beijing univ
- Chinese Academy of Sciences(NLPR-IACAS)
- Chinese Univ of Hong Kong + Hong Kong Polythechnic Univ+ Tsinghua Univ
- DAEDALUS, S.A.

- Hiroshima City Univ
- Information and Communications Univ
- Chinese Academy of Sciences(ISCAS)
- Keio Univ
- City Univ of Hong Kong
- National Taiwan Univ
- NEC
- Northeastern Univ
- Peking Univ
- Pohang Univ of Science and Technology
- Swedish Institute of Computer Science
- Technical Univ of Darmstadt
- Graduate Univ for Advanced
- Tornado Technologies Co., Ltd.,
- Toyohashi Univ of Technology
- Univ of Neuchatel
- Univ of Sussex

[Must]
- Hiroshima City Univ
- Keio Univ
- Mie Univ
- NICT
- NEC
- Ochanomizu Univ (2 Groups)
- Okayama Univ
- Osaka Prefecture Univ
- Otaru Univ of Commerce
- Tokyo Metropolitan Univ
- Tokyo Denki Univ
- Univ of Sheffield
- Yokohama National Univ

[PAT MIN]
- Hiroshima City Univ
- Hitachi, Ltd.,
- Huafan Univ
- Nagaoka Univ of Technology
- Northeastern Univ
- NTT Corporation
- Peking Univ
- Shenyang Institute of Aeronautical Engineering
- Toyohashi Univ of Technology
- Univ of California, Berkeley
- Univ of Montreal
- Xerox

[PAT MT]
- Fudan Univ
- Harbin Institute of Technology + Heilongjiang Institute of Technology
- Hitachi,Ltd.,
- Japan Patent Information Organization
- Kyoto Univ
- Massachusetts Institue of Technology
- Nara Institute of Science and Technology + NTT
- NICT
- National Taiwan Normal Univ
- NTT Corporation
- Pohang Univ of Science and Technology
- TOSHIBA
- Tottori Univ
- Toyohashi Univ of Technology + Hosei University
- Univ of Tsukuba

Types of Information Access

Exploratory Search

Look up

Learn

Investigate
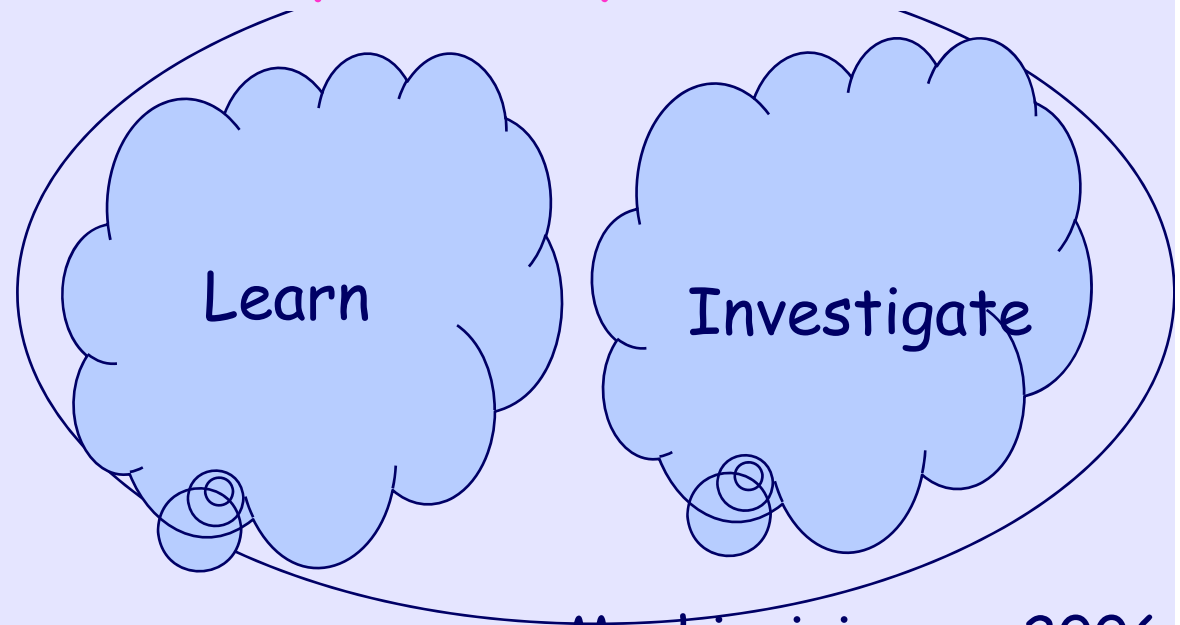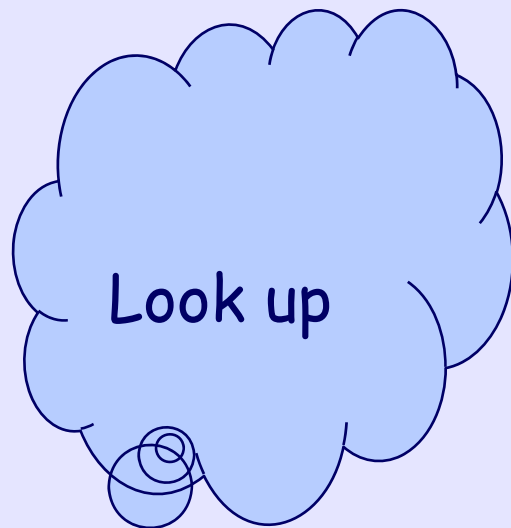
Machionini cacm 2006

Needs behind Queries
Human-Like Document Understanding

Call for NTCIR-8 task proposals

- Let's work together to construct a better infrastructure to encourage information-access research to move forward. Resources constructed in past NTCIRs are also available.

- Due to 30th November 2008
  - Write to Noriko Kando

For details, please click on each column in the table.

	Ad Hoc/ CLIR [Scientific Abstracts] (Japanese/ English IR))	Chinese IR	CLIR [News] (Cross-Lingual QA)	CLQA (Cross-Lingual IR)	MuST (Multimodal Summarization for Trend Information)	OPINION (Opinion Analysis)	PATENT	QAC (Question Answering)	TMREC (Term Recoginition)	TSC (Summa-rization)	WEB
NTCIR-1	NTCIR-1 Ad Hoc/ CLIR	-	-	-	-	-	-	-	NTCIR-1 TMREC	-	-
NTCIR-2	NTCIR-2 Ad Hoc/ CLIR	CIRB 010	-	-	-	-	-	-	-	NTCIR-2 SUMM	-
NTCIR-3	-	-	NTCIR-3 CLIR	-	-	-	NTCIR-3 PATENT	NTCIR-3 QA	-	NTCIR-3 SUMM	NTCIR-3 WEB
NTCIR-4	-	-	NTCIR-4 CLIR	-	-	-	NTCIR-4 PATENT	NTCIR-4 QA	-	NTCIR-4 SUMM	NTCIR-4 WEB
NTCIR-5	-	-	NTCIR-5 CLIR	NTCIR-5 CLQA	-	-	NTCIR-5 PATENT	NTCIR-5 QA	-	-	NTCIR-5 WEB
NTCIR-6	-	-	NTCIR-6 CLIR	NTCIR-6 CLQA	NTCIR-6 MuST	NTCIR-6 OPINION	NTCIR-6 PATENT	NTCIR-6 QA	-	-	-

NTCIR-7 PC Meeting@NTCIR-6

Mark Sanderson, Doug Oard, Atsushi Fujii, Tatsunori Mori, Fred Gey, Noriko Kando (and others)

Acknowledgments

- Japan Intellectual Property Association (JIPA)
- Industrial Property Cooperation Center, Japan
- Japan Parent Office
- Japan Patent Information Organization (JAPIO)
- Mainichi Newspaper
- NRI Cyber Patents
- PATOLIS
- Task organizers
- Participants and test-collections' users
- Information Retrieval Facility

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/