NTCIR-10 Conference & EVIA 2013

Noriko Kando, Tsuneaki Kato, Douglas W. Oard, Mark Sanderson
NTCIR-10 General Chairs

19th June 2013

NTCIR **N**II **T**estbeds and **C**ommunity for **I**nformation access **R**esearch

Research Infrastructure for Evaluating Information Access

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, forum

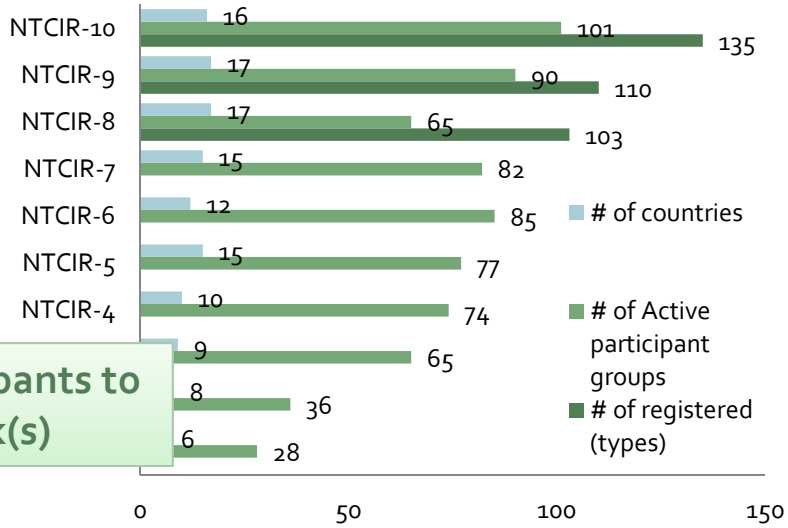
Project started in late 1997

- 18 months Cycle

2,347 Users!

Data sets (Test collections or TCs)

- Scientific, news, patents, web, CQA, Wiki, Exams
- Chinese, Korean, Japanese, and English



Participants to Task(s)

Tasks (Research Areas)

- IR: Cross-lingual tasks, patents, web, Geo, Spoken
- QA：Monolingual tasks, cross-lingual tasks
- Summarization, trend info., patent maps, Inference,
- Opinion analysis, text mining, Intent, Link Discovery, Visual

NTCIR-10 (2012-2013)
135 Teams Registered to Task(s)
973 Teams Registered so far

Tasks in the NTCIR

Domain	NTCIR th / Year	1 '99	2 '01	3 '02	4 '04	5 '05	6 '07	7 '08	8 '10	9 '11	10 '13	Tasks
Social Media・UGC									■			Community QA
							■	■	■			Opinion Analysis
Module-Based											■	IR + Inference
								■	■			QA + IR
Special Domain											■	MedNLP
											■	Math Retrieval
										■	■	Entrance Exam
										■	■	Spoken Doc Retireval
									■	■		GeoTime Retrieval
				■	■	■	■	□	□			Patent Retrieval
Quesiton Answering							■	■	■	□		Any types of Qs
					■	■						Dialogue
						■	■	■	■			Cross-Lingual IR
				■	■	■	■					Factoid, List
IE & IR										■	■	Link Discovery
Semantic										■	■	Inference
Summarization					■	■	■	■	■			Text Mining/
						■	■	■				Trend Info Visualization
			■	■	■							Text Summarization
Interactive										■		Interactive IR and Visualization
Web										■	■	Intent mining, Diversified Searc
										■	■	SERPT Quality
				■	■	■						Web Search
Cross-Lingual								■	■	■	■	Machine Translation
		■	■	■	■	■	■	■	■			Cross-Lingual IR
		■	■	■	■	■	■	■	■			Non-EnglishIR
Tetxt Retrieval		■	■	■	■	■	■	■	■			Ad Hoc, IR for QA

Year that the conference was held, The Tasks started 18 Months before

Tasks in NTCIR

Procedures in NTCIR Workshops

- Call for Task Proposals
- Selection of Task Proposals by Program Committee
- Discussion about Task Design in Each Task
- Registration to Task(s)
  - Deliver Training Data (Documents, Topics, Answers)
    - Experiments and Tuning by Each Participants
  - Deliver Test Data (Documents and Topics)
    - Experiments by Each Participants
- Submission of Experimental Results
- Pooling the Answer Candidates from the Submissions, and Conduct Manual Judgments
- Return Answers (Relevance Judgments) and Evaluation Results
- Conference Discussion for the Next Round
- Test Collection Release for non-participants

About NTCIR-10

What was new?

NTCIR-9 & 10: Objectives

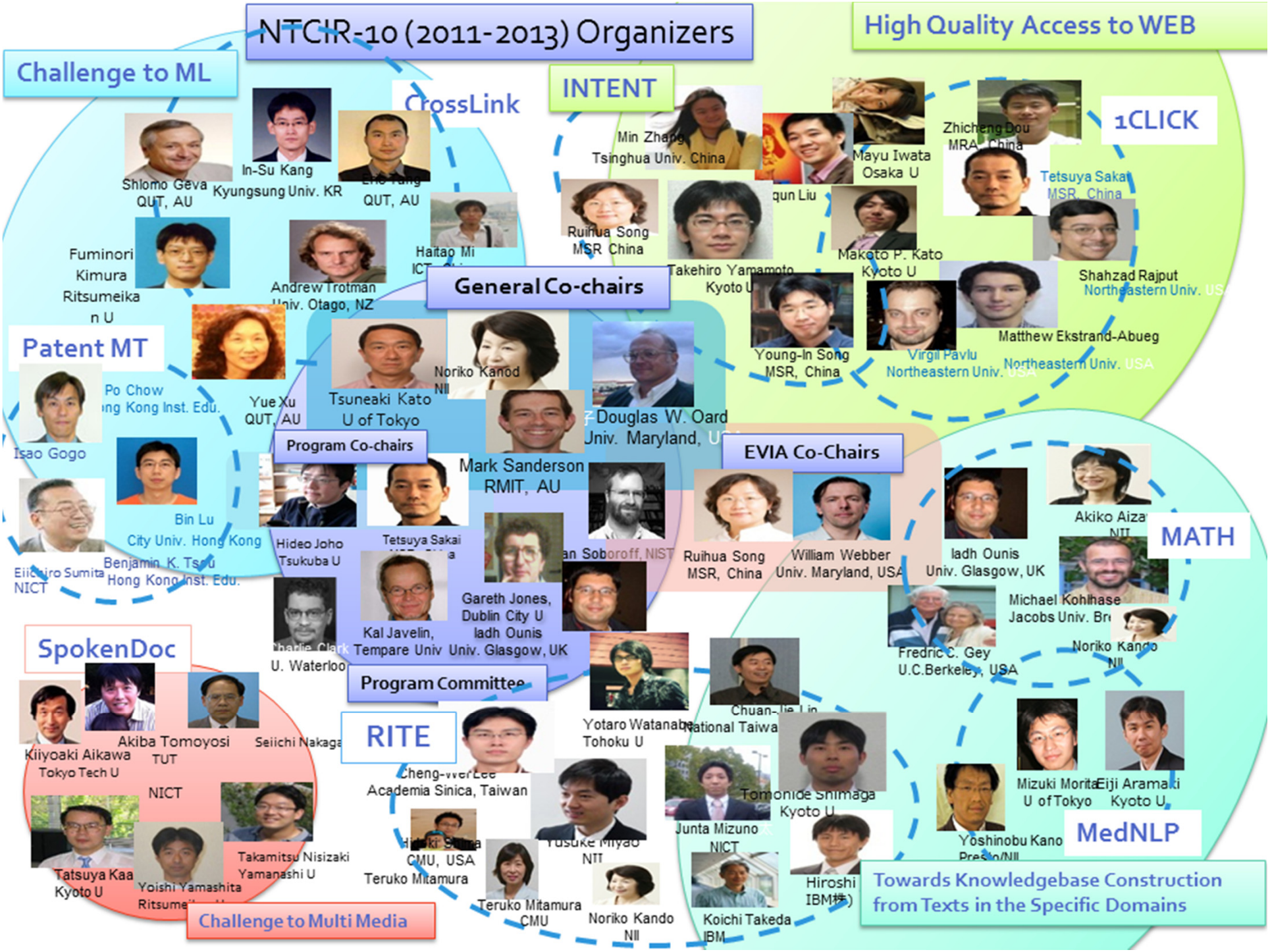
- Solid foundation
  - New structure
- Task diversity
  - Covers a wide context in Information Access
  - Studies rich media types
- Community-led task organisation
  - Sustainability of research
- Promotion of research resources
  - Show case in the NTCIR-9 Meeting

NTCIR-10: Structure

- General Co-Chairs
  - Noriko Kando (NII)
  - Tsuneaki Kato (The Univ of Tokyo)
  - Douglas W. Oard(Univ of Maryland)
  - Mark Sanderson (RMIT Univ.)
- Program Co-Chairs
  - Hideo Joho (Tsukuba University)
  - Tetsuya Sakai (MSR)
- Task Organisers
  - 58 researchers worldwide
  - 135 Groups Registered, 101 Active Participating Groups
- EVIA Co-Chairs
  - Ruihua Song (MSR)
  - William Webber (Univ of Maryland)

Highest Numbers!

Highest Numbers!


NTCIR-10 (2011-2013) Organizers
Challenge to ML
CrossLink
INTENT
High Quality Access to WEB
1CLICK
Shlomo Geva
QUT, AU
In-Su Kang
Kyungsung Univ. KR
Eric Yang
QUT, AU
Min Zhang
Tsinghua Univ. China
Mayu Iwata
Osaka U
Zhicheng Dou
MRA, China
Tetsuya Sakai
MSR, China
Ruihua Song
MSR China
Takehiro Yamamoto
Kyoto U
Makoto P. Kato
Kyoto U
Shahzad Rajput
Northeastern Univ. US
Fuminori Kimura
Ritsumeikan U
Andrew Trotman
Univ. Otago, NZ
Haitao Mi
ICT
General Co-chairs
Patent MT
Noriko Kanod
NII
Tsuneaki Kato
U of Tokyo
Douglas W. Oard
Univ. Maryland, US
Young-In Song
MSR, China
Virgil Pavlu
Northeastern Univ. USA
Matthew Ekstrand-Abueg
Northeastern Univ. USA
Po Chow
Hong Kong Inst. Edu.
Yue Xu
QUT, AU
Isao Gogo
Program Co-chairs
EVIA Co-Chairs
Mark Sanderson
RMIT, AU
Bin Lu
City Univ. Hong Kong
Hideo Joho
Tsukuba U
Tetsuya Sakai
Benjamin K. Tsou
Hong Kong Inst. Edu.
Eiichiro Sumita
NICT
Ruihua Song
MSR, China
William Webber
Univ. Maryland, USA
Iadh Ounis
Univ. Glasgow, UK
Akiko Aizawa
NII
MATH
Michael Kohlhase
Jacobs Univ. Bremen
Noriko Kando
NII
Fredric C. Gey
U.C.Berkeley, USA
Gareth Jones,
Dublin City U
Iadh Ounis
Univ. Glasgow, UK
Kal Javelin,
Tempare Univ
Charlie Clark
U. Waterloo
SpokenDoc
Program Committee
Akiba Tomoyosi
TUT
Seiichi Nakagawa
Kiiyoaki Aikawa
Tokyo Tech U
NICT
RITE
Yotaro Watanabe
Tohoku U
Chuan-Jie Lin
National Taiwan
Cheng-Wei Lee
Academia Sinica, Taiwan
Tomonide Shimaga
Kyoto U
Junta Mizuno
NICT
Mizuki Morita
U of Tokyo
Eiji Aramaki
Kyoto U
Yusuke Miyao
NII
Yoshinobu Kano
Preston/NII
MedNLP
Takamitsu Nisizaki
Yamanashi U
Tatsuya Kaa
Kyoto U
Yoichi Yamashita
Ritsume
Hiroshi
IBM
Teruko Mitamura
CMU, USA
Teruko Mitamura
CMU
Noriko Kando
NII
Koichi Takeda
IBM
Towards Knowledgebase Construction from Texts in the Specific Domains
Challenge to Multi Media

NTCIR-10 Tasks

Core

- [CrossLink-2] Cross-lingual Link Discovery
- [Intent-2] Search intent and diversification
- [1Click-2] One-Click Access
- [PatentMT-2] Cross-lingual access to Patent Docs
- [RITE-2] Recognizing Inference in Text
- [SpokenDoc-2] IR for Spoken Documents

Pilot ** NEW!! ***

- [Math] Access to mathematical contents
- [MedNLP] MedNLP

NTCIR-10 Program Committee

- Charles Clarke
- Kalervo Järvelin
- Hideo Joho, chair
- Gareth Jones
- Noriko Kando
- Tsuneaki Kato
- Tetsuya Sakai, chair
- Mark Sanderson
- Ian Soboroff

NTCIR-10 Data Provider/Cooperation

CMU (Clue Web 09)

Hong Kong Institute of Education

Japan Patent Office

Linguistic Data Consortium

Mainichi Newspaper

Microsoft Research

National Center for University Admissions

National Institute of Information and Communication Technology (NICT)

National Institute for Japanese Language and Linguistics (NIJL)

Sogou.com

Spoken Document Processing WG, SIG Spoken Language Processing, Information Processing Society Japan

Tokyo Shoseki

UDN.COM

Zhuhai Chilinstar

And all the organizations of the task organizers !

NTCIR

NTCIR-10 Conference

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

12

NTCIR-10 Conference

Highest!
(+16)

- Conference!
- 218 Pre-Registered Participants!
- Keynote: Shashi Thakker (Google)
  Charles Clarke (U of Waterloo)
- Invited talks:
  - Ian Soboroff (TREC/TAC)
  - Gareth Jones (MedEval)

NTCIR-10 Conference Organization Committee

General Co-Chairs:

Noriko Kando, Tsuneaki Kato, Douglas W. Oard, Mark Sanderson

Accounting Chair : Hiroaki Ohshima

Exhibitions and Poster Chair : Masaharu Yoshioka

Local Arrangements Chair : Emi Ishita

Publications Chair : Kazuaki Kishida, Kazuko Kuriyama

Sponsor Chair : Tsuneaki Kato

Travel Support Chair : Takehito Utsuro

Special Thanks to

NTCIR-10 Conference Sponsors

Silver

10 Travel Supports!

Hitachi, Ltd.

NEC Information Systems, Ltd.

Bronze

Japan Patent Information Organization (Japio)

IR-Advanced Language Technologies, Inc.

Yahoo Japan Corporation

A9.Com / Amazon Japan

NTT Data Corporation

NTCIR-10

Assessors

NTCIR-10

NTCIR Project Office

Thanks !

Mariko Okada

Kazuko Matsuo

Noriko Kubo

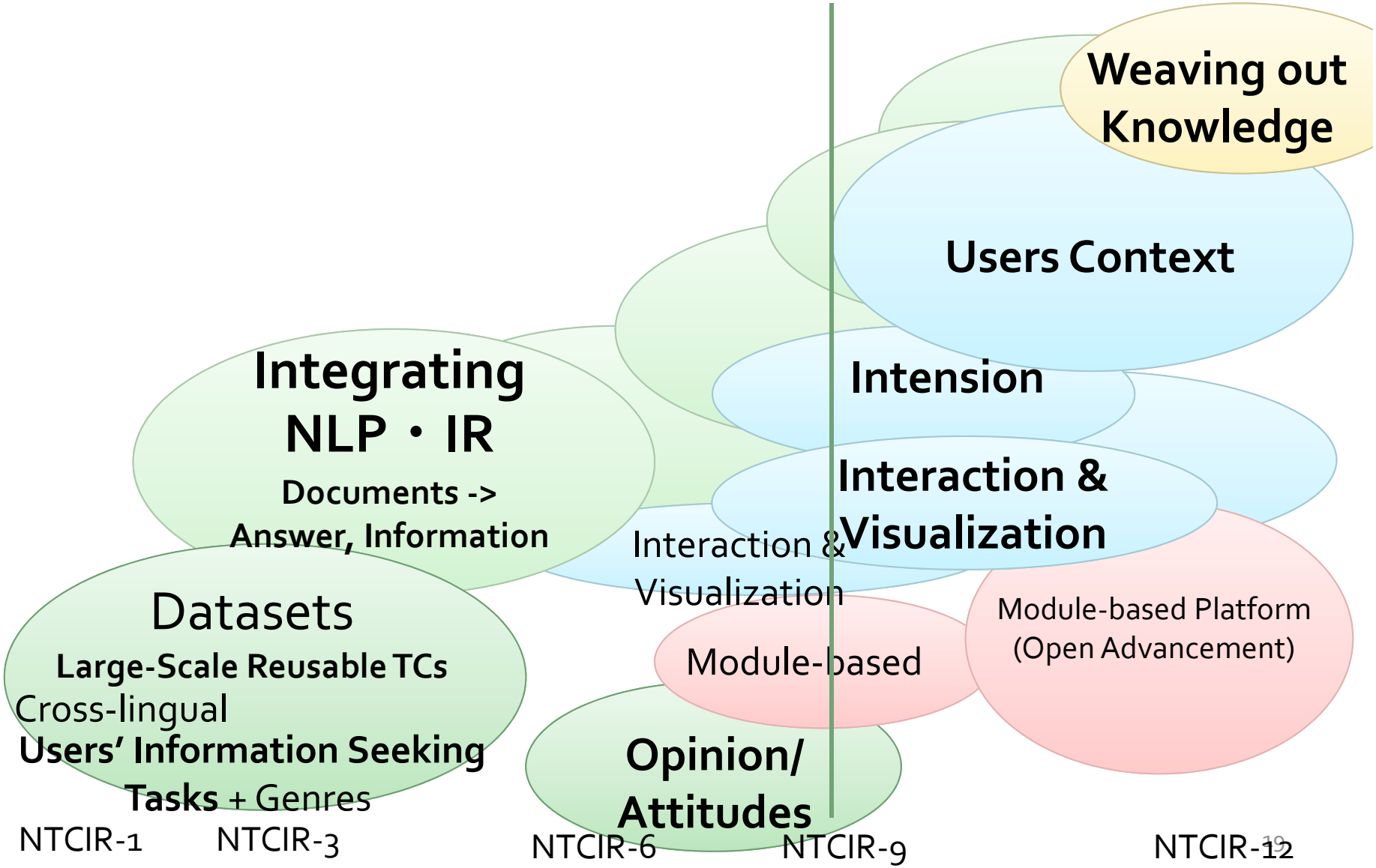
Miho Sugimoto (by June 2012)

NII Testbeds and Community for Information access Research

Why?

-Data sets / Users' Information Seeking Tasks
-Evaluation Methodology
-Reusable vs Reproducibility
-User-Centered Evaluation
-Experimental Platforms
-Open Advancement
-Advanced NLP → Knowledge- or Semantic-based
-Diversified IA Applications in the Real World
-Best Practice for a technology
→ Best Practice for Evaluation Methodology
-Big Data (Documents + Behaviour data)

Mid-term Vision



http://research.nii.ac.jp/ntcir/ntcir-10/

Thank you for your participation! Enjoy!

For further enquiries, contact the NTCIR office
ntc-secretariat@nii.ac.jp