

# Overview of NTCIR

Charlie Clarke, Noriko Kando  
NTCIR-14 General Chairs

■ A series of evaluation workshops designed to enhance research in information-access technologies by providing an infrastructure for evaluations. ■ Data sets, evaluation methodologies, forum

### Project started in late 1997

■ 18 months Cycle

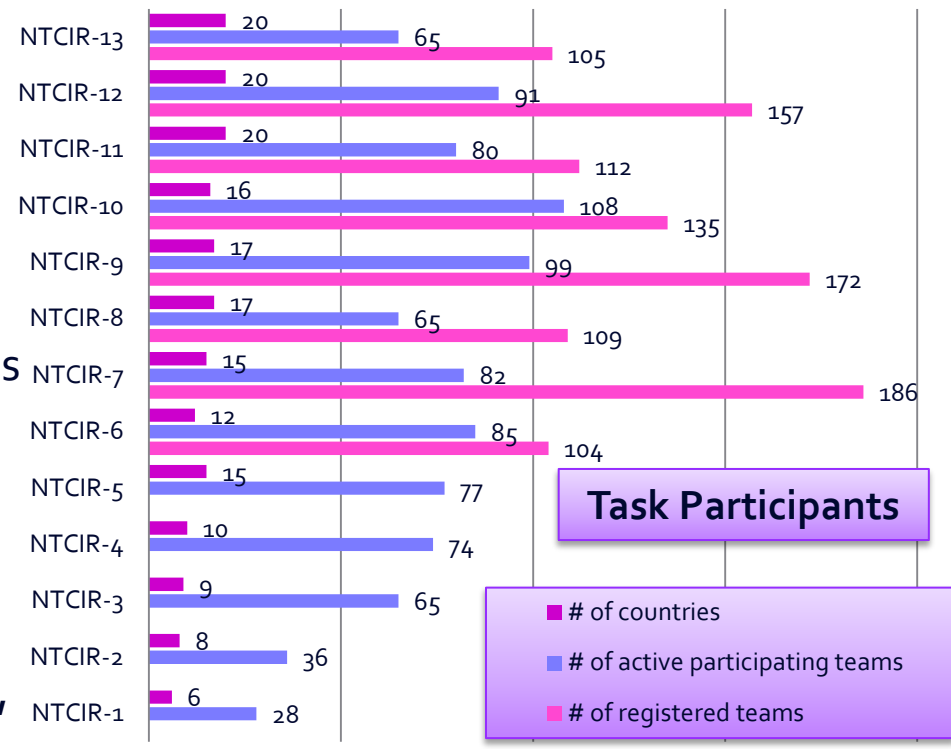
### Data sets (Test collections or TCs)

**4146 Users!**

- Scientific, news, patents, web, CQA, Wikipedia, exams, spoken, microblogs, lifelogs
- Chinese, Korean, Japanese, and English

### Tasks (Research Areas)

- IR: cross-lingual, patents, web, geo, time, math, Spoken, recipe, microblogs, lifelogs
- QA : monolingual, cross-lingual, interactive, IR for QA, complex, exams
- summarization, trend info., patent maps, Inference, 2-layer summarization for mobile
- opinion analysis, text mining, Intent, link discovery, Visual, info extraction,



# Tasks in NTCIRs

NTCIR th	1	2	3	4	5	6	7	8	9	10	11	12	Tasks		
Domain Year	'99	'01	'02	'04	'05	'07	'08	'10	'11	'13	'14	'16			
Social Media • UGC								■				■	Microblog		
						■	■	■					Community QA		
													Opinion Analysis		
Module-Based							■	■		■			IR + Inference		
													QA + IR		
Special Domain													■	Lifelog images	
												■		Recepe	
											■	■	■	MedNLP	
											■	■	■	Math Retrieval	
										■	■	■	■	Exam (Entailment, QA)	
										□	□	■	■	Spoken Q& Doc Retrieval	
									■	■				■	Temporalia
				■	■	■	■	□	□						GeoTime Retrieval
Question Answering													■	■	Exam (multi-choice, essay)
							■	■	■	□					Any types of Qs
				■	■										Dialogue
					■	■	■	■	■						Cross-Lingual IR
			■	■	■	■									Factoid, List
IE & IR									■	■					Link Discovery
Semantic									■	■	■				Inference
Summarization												■	■		2-layer (mobile)
				■	■	■	■	■							Text Mining/
		■	■	■		■	■	■							Trend Info Visualization
Interactive															Text Summarization
													■	■	Micoblog
Mobile												■	■		Spoken Query
									■						Visualization
Web															MobileClick
										■	■	■	■		Intent, Diversification
				■	■	■				■	■				SERPT Quality
Cross-Lingual															Web Search
	■	■	■	■	■	■	■	■	■	■					Machine Translation
	■	■	■	■	■	■	■	■							Cross-Lingual IR
Text Retrieval															Non-EnglishIR
	■	■	■	■	■	■	■	■							Ad Hoc, IR for QA

# Tasks in NTCIRs

Year	1999	2001	2002	2004	2005	2007	2008	2010	2011	2013	2014	2016	2017
<b>Task/NTCIR round</b>	1	2	3	4	5	6	7	8	9	10	11	12	13
<b>Total number</b>	37	39	61	74	79	81	80	66	102	108	93	97	71
Automatic Term Recognition and Role Analysis (TMREC) (1)	9												
Ad hoc/Crosslingual IR (1) → Chinese/English/Japanese IR (2) → CLIR (3-6)	28	30	20	26	25	22							
Text Summarization Challenge (TSC) (2-4)		9	8	9									
Web Retrieval (WEB) (3-5)			7	11	7								
Question Answering Challenge (QAC) (3-6)			16	18	7	8							
Patent Retrieval [and Classification] (PATENT) (3-6)			10	10	13	12							
Multimodal Summarization for Trend Information (MUST) (5-7)					13	15	13						
Crosslingual Question Answering (CLQA) (5, 6) → Advanced Crosslingual Information Access (ACLIA) (7, 8)					14	12	19	14					
Opinion (6) → Multilingual Opinion Analysis (MOAT) (7, 8)						12	21	16					
Patent Mining (PAT-MN) (7, 8)							12	11					
Community Question Answering (CQA) (8)								4					
Geotemporal IR (GeoTime) (8, 9)								13	12				
Interactive Visual Exploration (Vis-Ex) (9)									4				
Patent Translation (PAT-MT)(7, 8) → Patent Machine Translation (PatentMT)(9, 10)							15	8	21	21			
Crosslingual Link Discovery (Crosslink) (9, 10)									11	10			
INTENT(9, 10) → Search Intent and Task Mining (IMine) (11, 12)									16	11	12	9	
One Click Access (1CLICK)(9, 10) → Mobile Information Access (MobileClick) (11, 12)									4	8	4	11	
Recognizing Inference in Text (RITE)(9,10) → Recognizing Inference in Text and Validation (RITE-VAL)(11)									24	28	23		
IR for Spoken Documents (SpokenDoc) (9, 10) → Spoken Query and Spoken Document Retrieval (SpokenQuery&Doc) (11, 12)									10	12	11	7	
Mathematical Information Access (Math) (10, 11) → MathIR (12)										6	8	6	
Medical Natural Language Processing (MedNLP) (10, 11) → MedNLPDoc (12) → MedWeb (13)										12	12	8	9
QA Lab for Entrance Exam (QALab) (11, 12, 13)											11	12	11
Temporal Information Access (Temporalia) (11, 12)											8	14	
Cooking Recipe Search (RecipeSearch) (11)											4		
Personal Lifelog Organisation & Retrieval (Lifelog) (12, 13)												8	4
Short Text Conversation (STC) (12, 13)												22	27
Open Live Test for Question Retrieval (OpenLiveQ) (13)													7
Actionable Knowledge Graph (AKG) (13)													3
Emotion Cause Analysis (ECA) (13)													3
Neurally Augmented Image Labelling Strategies (NAILS) (13)													2
We Want Web (WWW) (13)													5

MedWeb  
QA Lab

Lifelog  
STC

AKG

ECA

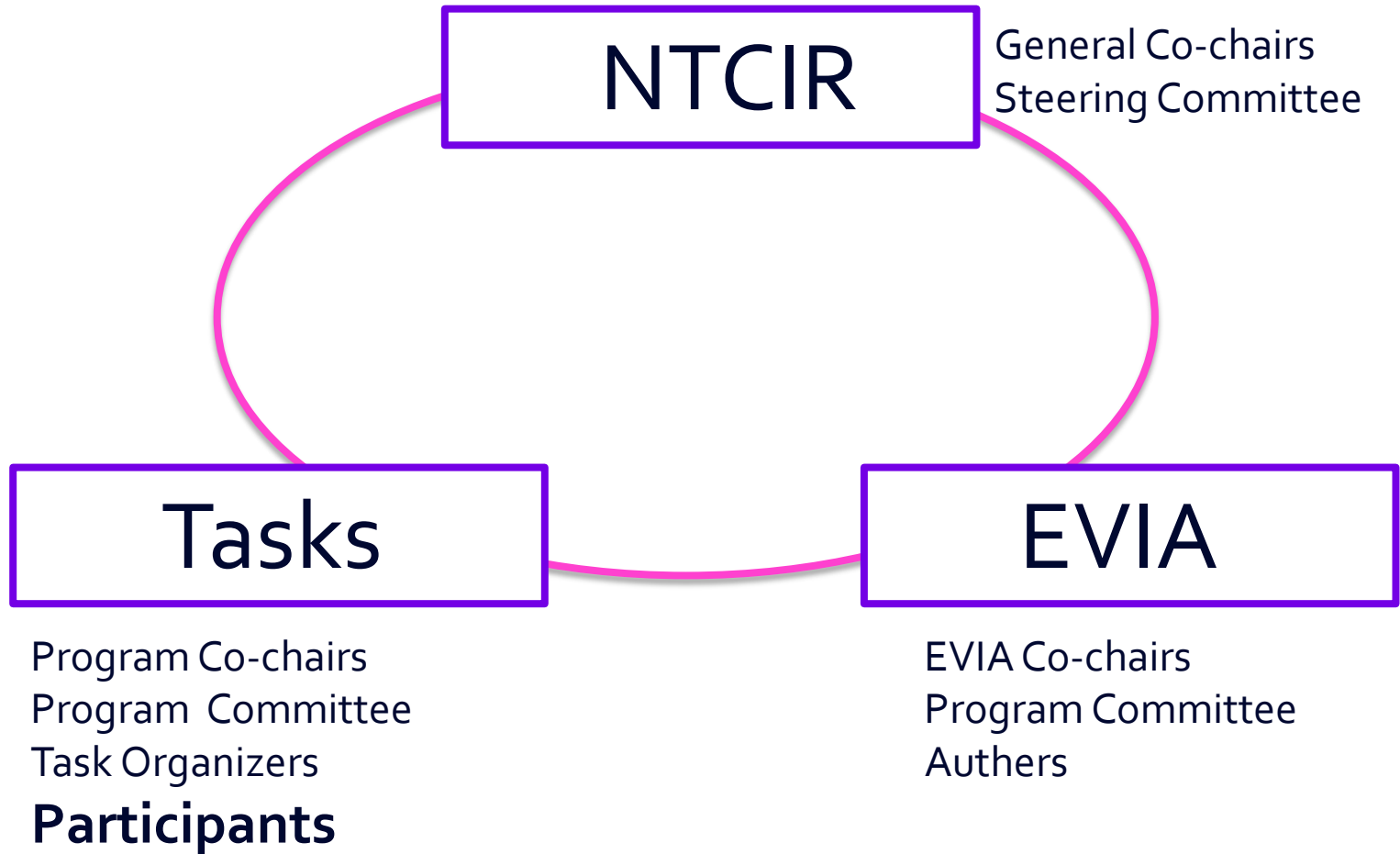
NAILS

WWW

# IR Evaluations

- TREC: Text Retrieval Conference
- CLEF: Cross-Language Evaluation Forum  
→ Conference and Labs of the Evaluation Forum
- NTCIR: NII Testbeds and Community for Information access Research

# Structure of NTCIR



# NTCIR changed name 3 times

- Nacsis Test Collection for IR research
  - NTCIR-1
- NII Test Collection for IR research
  - NTCIR-2~8
- NII Testbeds and Community for Information access Research
  - NTCIR-9
- NTCIR is a **Community-Lead Activities** for **Creating Future Values**  
so, Let's work Together!

# Procedures in NTCIRs

- Call for Task Proposals
  - Selection of Task Proposals by Program Committee
  - Discussion about Task Design in Each Task
- 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
  - Analysis of the results, and improving the systems, etc.
  - Paper writing for the Conference proceedings
- Conference
  - Presenting the achievements of this Round
  - Discussion for the Next Round
  - Networking
- Test Collection Release for non-participants



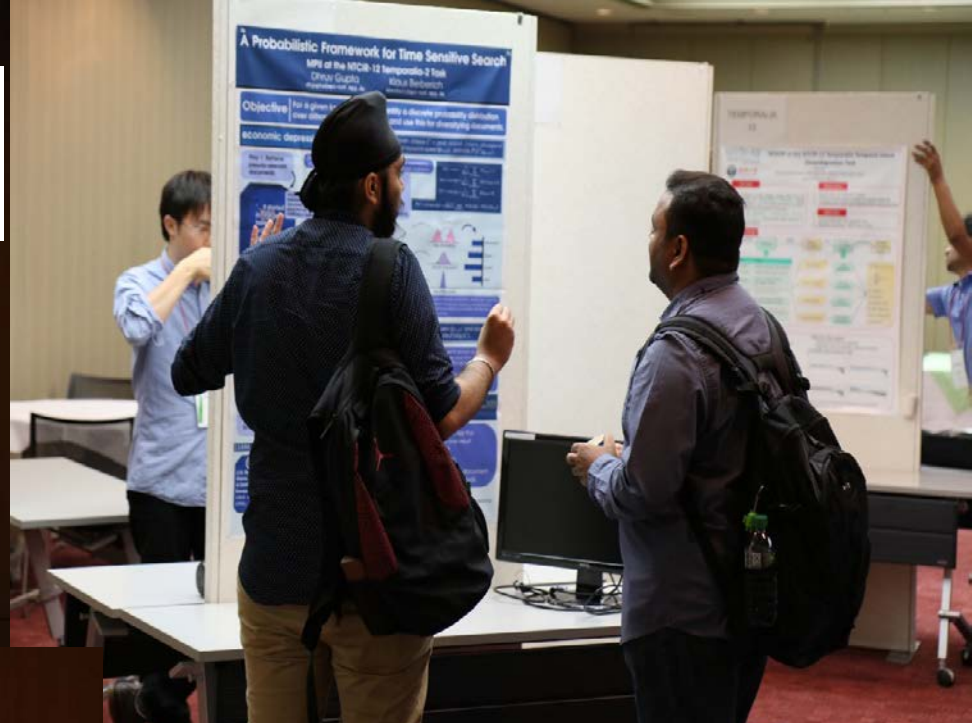
Kickoff



Pls join  
Discussion!!



# Past NTCIR Conference & EVIA



## Past NTCIR Conference & EVIA



## EVIA has attracted top experts on IR Evaluation

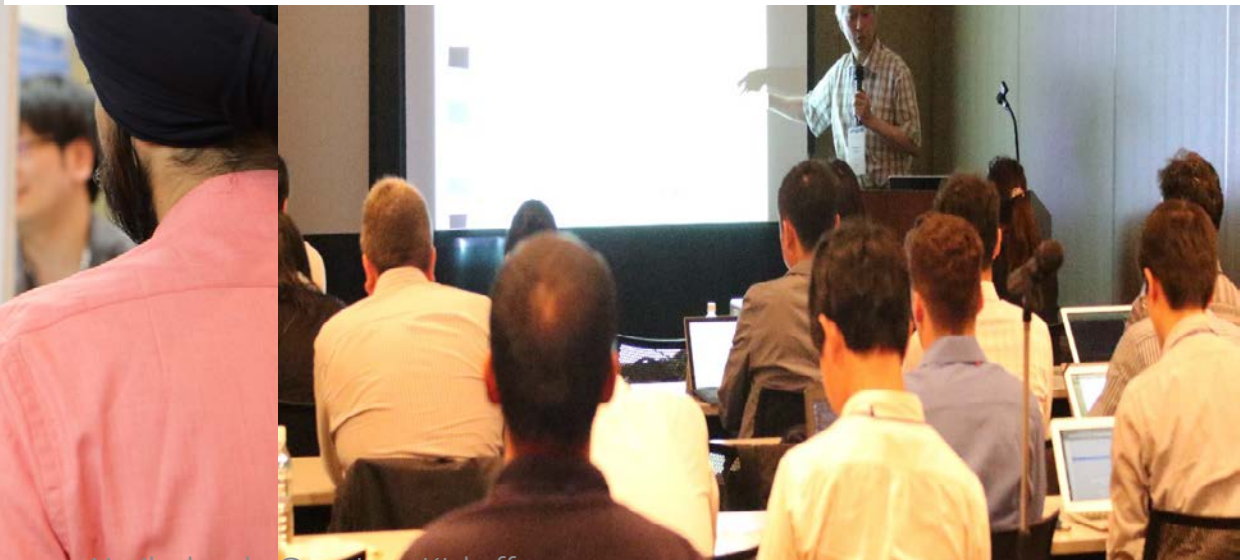






Past NTCIR Conference

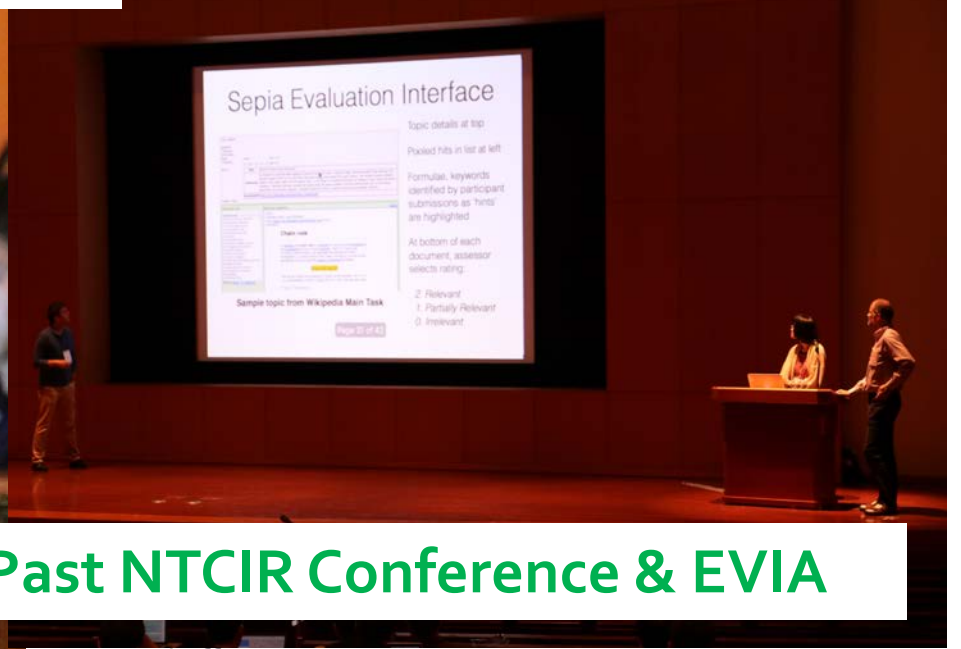
## Satelite Lifelog Workshop in Glasgow at NTCIR-12 Conference, June 2012



Noriko kando @ ntcir-14 Kickoff



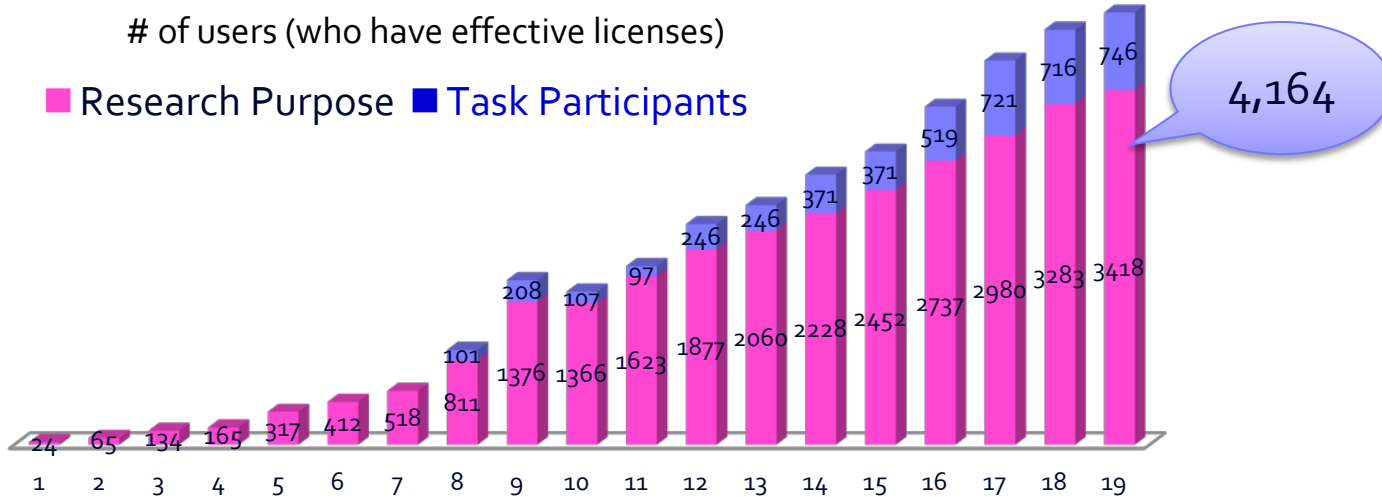
Travel Supports are provided for Overseas Task/EVIA Young Authors



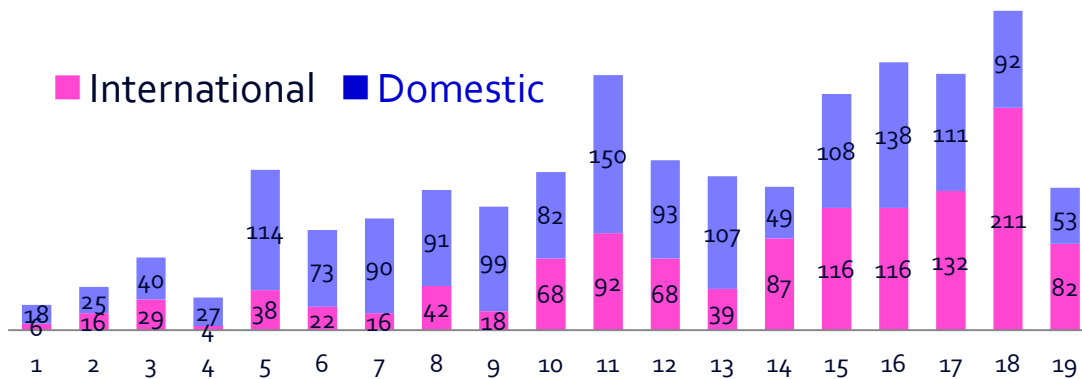
Past NTCIR Conference & EVIA



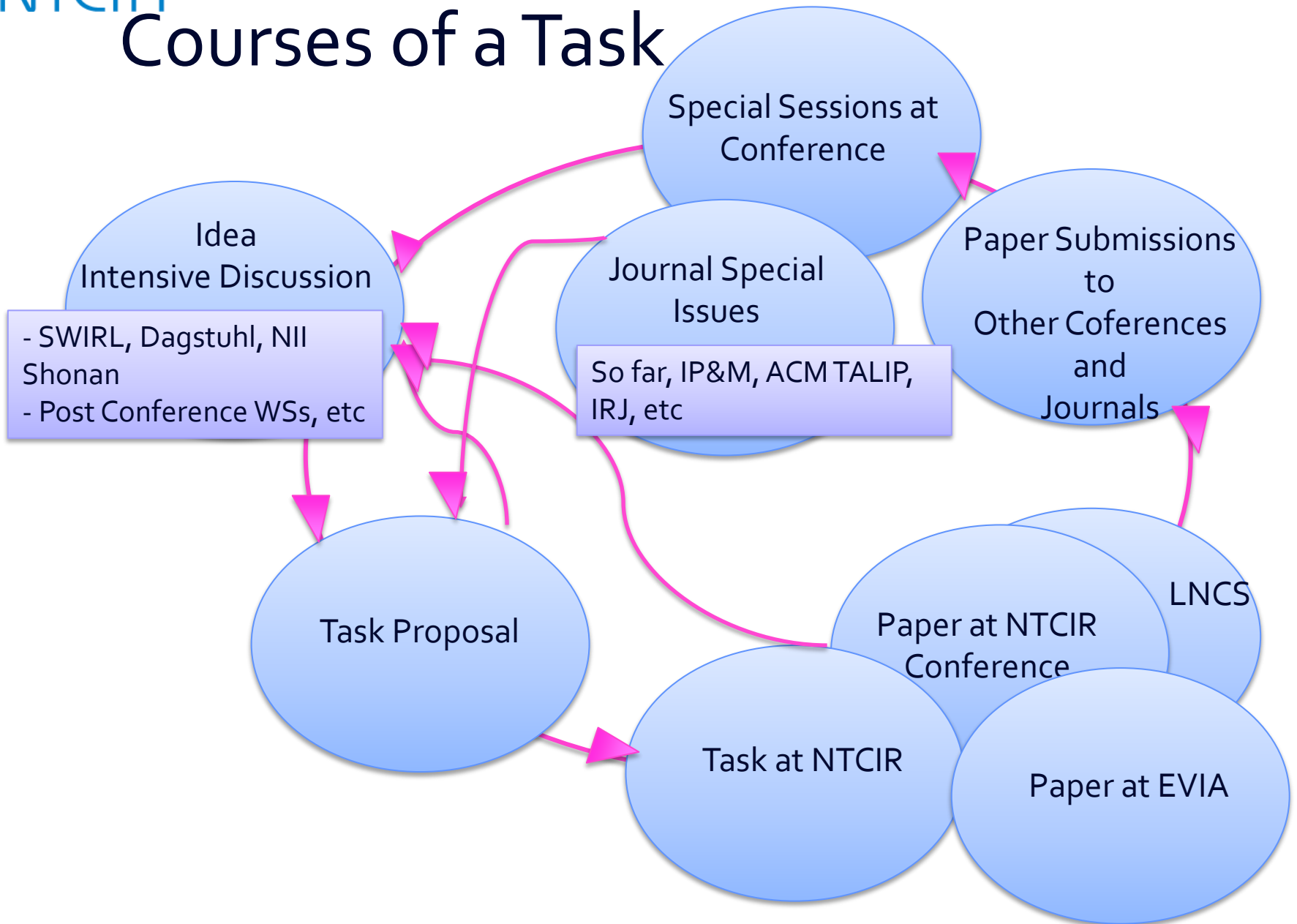
# Uses of the Test Collections by Community



## International Users vs Domestic Users



# Courses of a Task



# 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

