



**Welcome!**

# Overview of NTCIR

Noriko Kando, Charlie Clarke  
NTCIR-15 General Chairs

September 30<sup>th</sup> 2019

■ 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

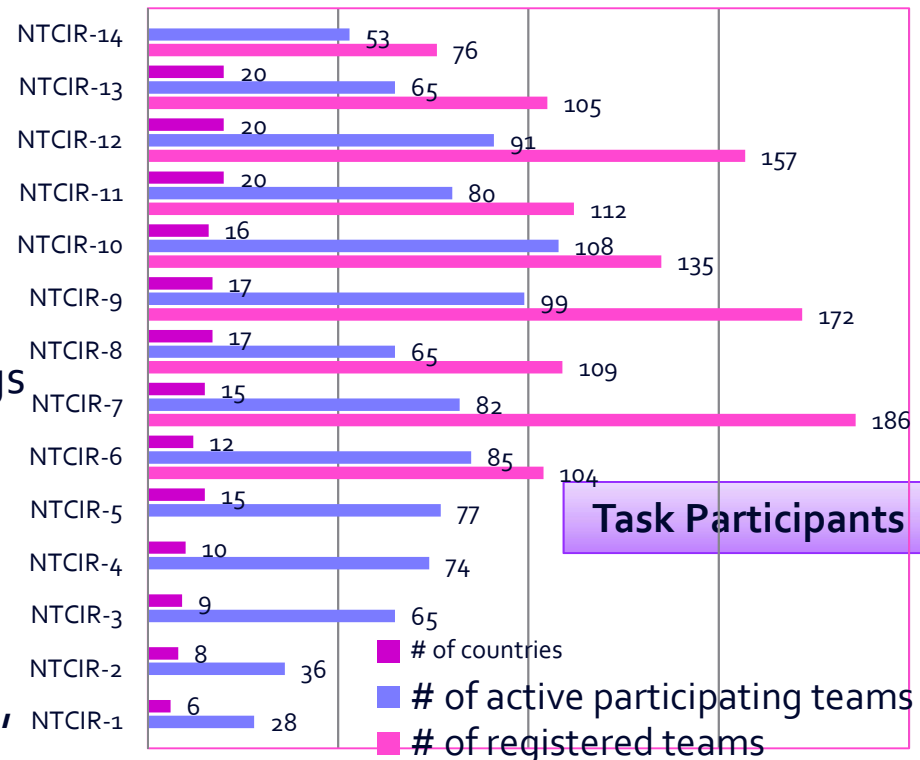
4,654 Users!

### Data sets (Test collections or TCs)

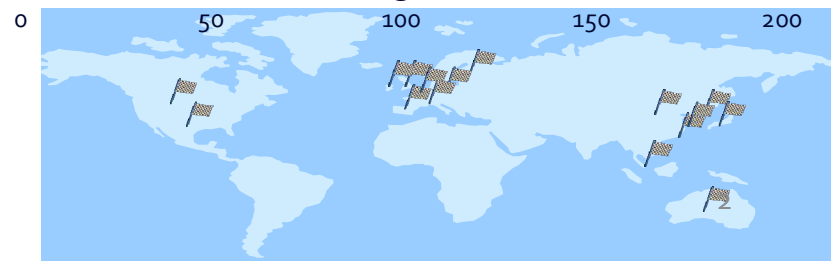
- Scientific, news, patents, web, CQA, Wikipedia, exams, spoken, microblogs, lifelogs
- Chinese, Korean, Japanese, and English

### Tasks (Research Areas)

- IR: cross-lingual, patents, web, geo, time, math, Spoken, recepe, 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,



Task Participants



# 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
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-English IR
	■	■	■	■	■	■	■	■					■ Ad Hoc, IR for QA

## Tasks in NTCIRs

Year	1999	2001	2002	2004	2005	2007	2008	2010	2011	2013	2014	2016	2017	2019
Task/NTCIR round	1	2	3	4	5	6	7	8	9	10	11	12	13	14
Total number	37	39	61	74	79	81	80	66	102	108	93	97	71	47
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)										12	12	8	9	
QA Lab for Entrance Exam (QALab) (11, 12, 13) → QA Lab for Political Information (QALab-PoliInfo) (14)											11	12	11	13
Temporal Information Access (Temporalia) (11, 12)											8	14		
Cooking Recipe Search (RecipeSearch) (11)											4			
Personal Lifelog Organisation & Retrieval (Lifelog) (12, 13, 14)												8	4	6
Short Text Conversation (STC) (12, 13, 14)												22	27	13
Open Live Test for Question Retrieval (OpenLiveQ) (13, 14)													7	4
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, 14)													5	4
Fine-Grained Numeral Understanding in Financial Tweet (FinNum) (14)														6
CLIR/NTCIR/TREC/REproducibility (CENTRE) (14)														1

QA Lab PoliINFO

Lifelog

STC

OpenLiveQ

WWW

FinNum

CENTRE

# About NTCIR-14

What was new?

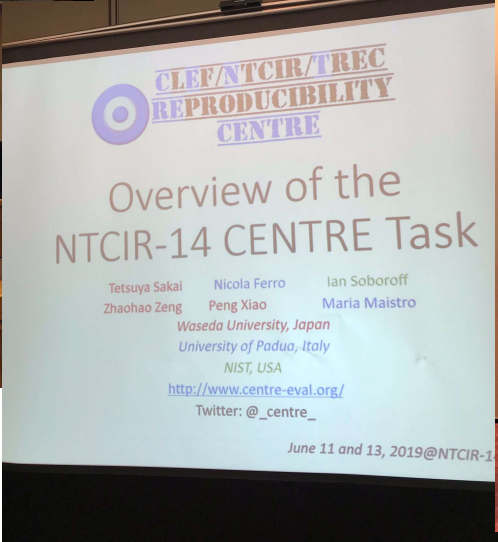
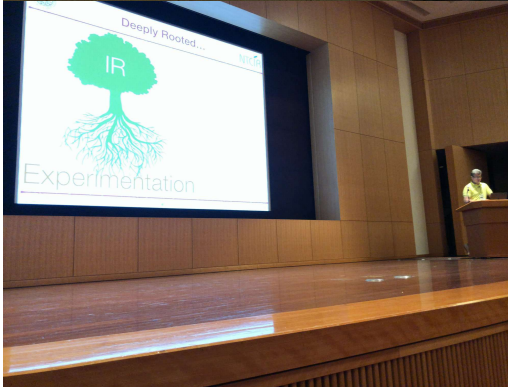
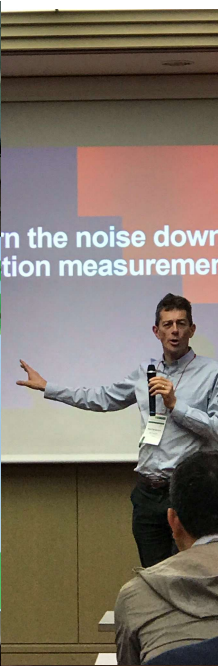
# NTCIR-14 Tasks

- **CORE**

- Lifelog Search Task ("Lifelog-3")
- Open Live Test for Question Retrieval ("OpenLiveQ-2")
- QA Lab for Political Information ("QALab PoliINFO") **NEW**
- Short Text Conversation ("STC-3")
- We Want Web ("WWW-2")

- **PILOT**

- CLEF/NTCIR/TREC Reproducibility ("CENTRE")
- Fine-Grained Numeral Understanding in Financial Tweet ("FinNum") **NEW**



# New Trial at NTCIR-14: LNCS Publication

- **Post-conference proceedings published by LNCS**
  - All the research groups that participated in NTCIR-14 were invited to submit a paper
  - At least 40% of the submitted paper should be a new content
  - The submitted papers were reviewed by the NTCIR-14 post-conference proceedings program committee members
- **Around 15 papers were accepted**



# NTCIR-14 Data & Infrastructure Providers

Thank you indeed!

- Yahoo! JAPAN Corporation (OpenLiveQ)
- ADAPT Centre and Dublin City University (Lifelog)
- JSPS Kakenhi 16H02912 and NII's Joint Research grant (QA Lab PoliINFO).
- Sogou, Inc (WWW)
- Carnegie Mellon University (WWW)
- MOST Joint Research Center for AI Technology (FinNum)
- All Vista Healthcare, Taiwan (FinNum)

# NTCIR-14 Conference

- Participants: 138 (as of June 3<sup>rd</sup>)
- **Keynote: Charles Clarke (U Waterloo)**
- Invited talks (DAY-2):
  - Nicola Ferro (Evaluation Measure Scale)
  - Gareth Jones (MediEval)
- Panel on Future IA Evaluation (Day-4)
- **EVIA Keynote (Day-1) Mark Sanderson**
- NTCIR Book Session (Day-1)
  - Doug Oard, Tetsuya Sakai
- CLEF Book/CLEF (Day-1)
  - Nicola Ferro

# NTCIR-14 Conference Organizing Committee

## **General Co-Chairs :**

Charlie Clarke, Noriko Kando

## **Local Arrangements :**

Tomoko Osuga

**Posters :** Hideyuki Shibuki

**Proceedings (Online) :** Makoto P Kato, Emi Ishita

**Proceedings (Springer LNCS) :** Yiqun Liu **NEW!**

**Web:** Madoka Ishioroshi

**Secretaries:** Hiroko Shoji, Hisayo Mizuno

# NTCIR-14 Conference Sponsors

**Special  
Thanks to**

Silver

Yahoo! Japan

10 Int'l Travel

Bronze

Japan Patent Information Organization (Japio)  
IR-Advanced Language Technologies, Inc. (IR-ALT)

Hitachi, Co.

Support

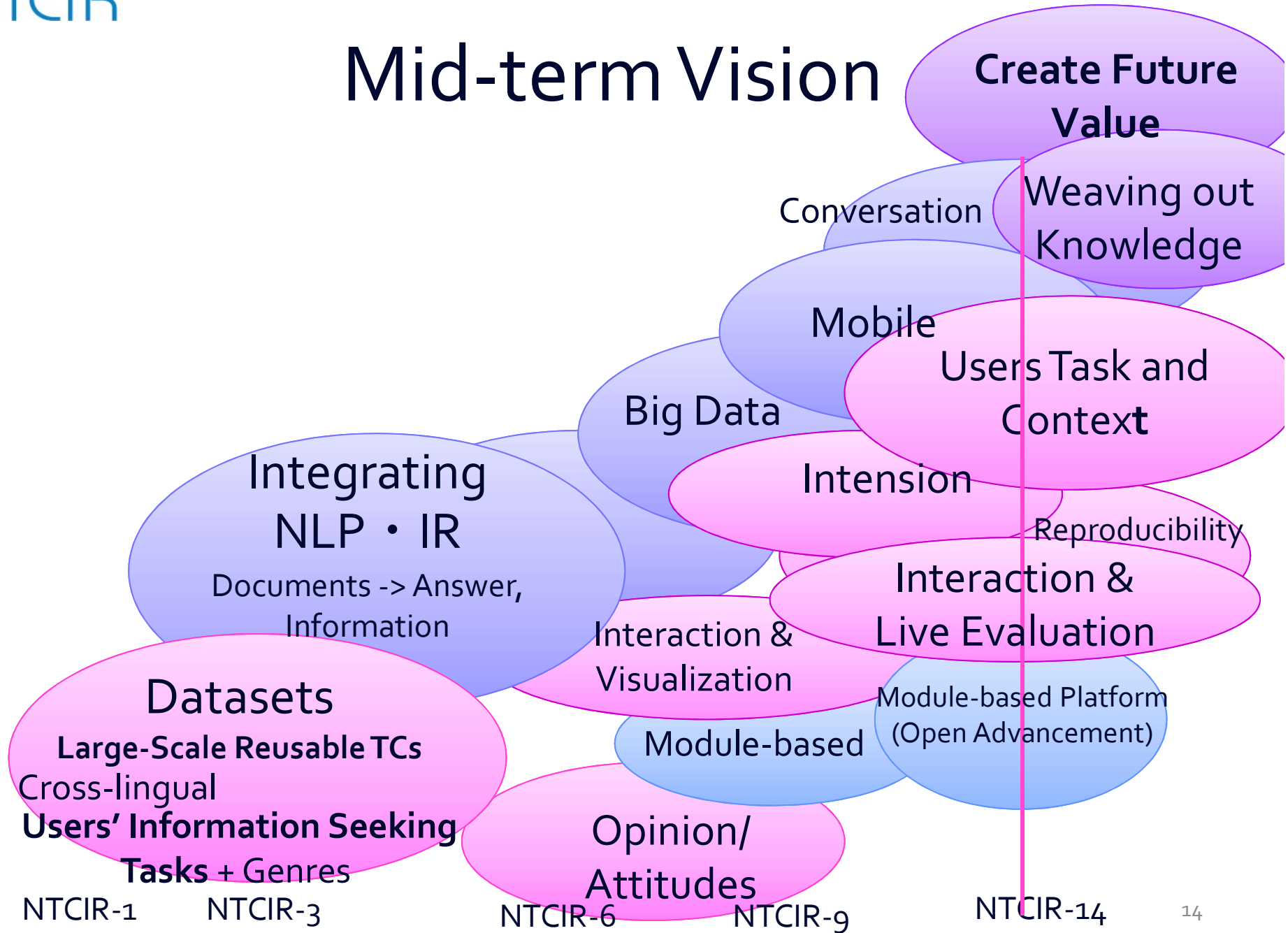
NICT

# 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



NTCIR

NTCIR

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

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