NTCIR

Proposal for NTCIR-11 Pilot task
QA Lab

Madoka Ishioroshi (NII)
Kelly Itakura (NII)
Noriko Kando (NII)
Yoshinobu Kano (PRESTO/NII)
Teruko Mitamura (CMU)
Hideyuki Shibuki (Yokohama National University)
Advisor: Tatsunori Mori (Yokohama National University)

NTCIR-11 introduction session at NTCIR-10 Conference
June 21, 2013

Thanks Madoka for slides

NTCIR

Road map

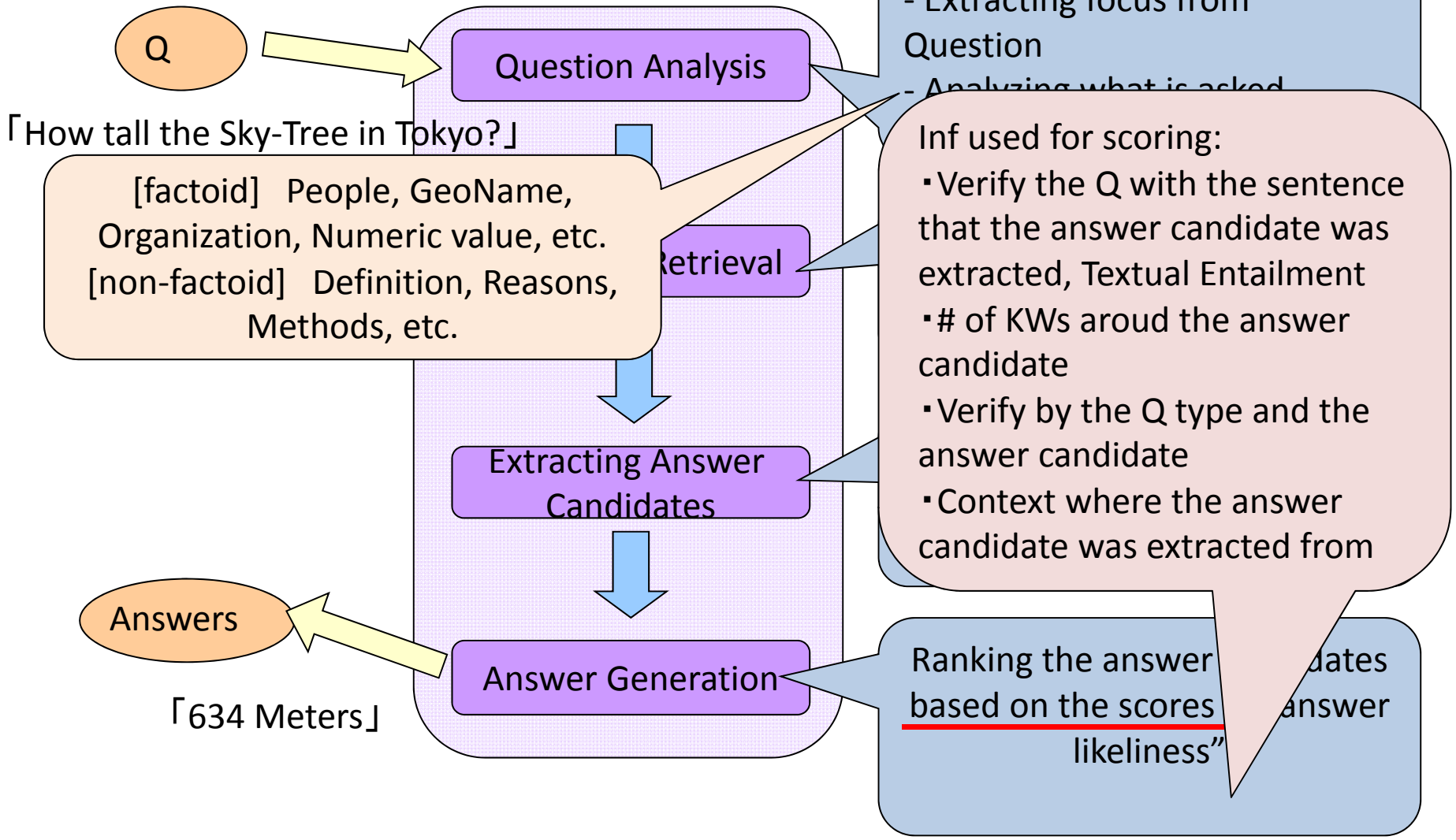
- Main Idea
- Background
  - QA
  - ACLIA: “Dream” system superior than the top end-to-end, but didn’t know why.
- Evaluation Method
  - Systems
  - What evaluate/ Metrics
  - Corpus
  - Resources

NTCIR

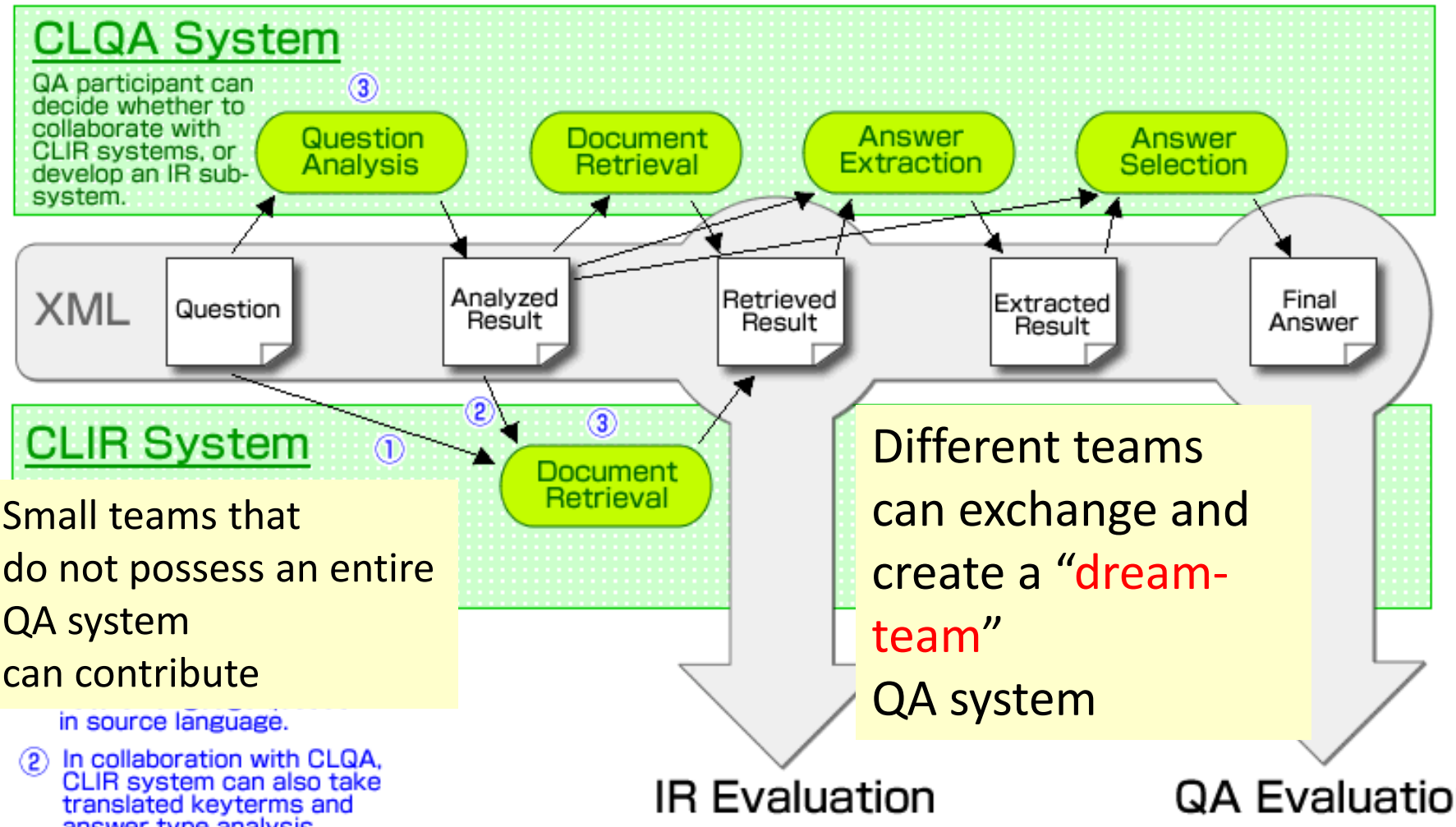
Main Idea

- Challenge for Real World Problem of QA
  - University Entrance Exams - New Q-types not covered by existing QA systems
- Challenge for a new structure of evaluation – Module-based, building workable system as a joint effort of all the participants
  - QA is very complicated
  - Hard for a research group to be good on every components
  - Your module used and rated by other people
- Continuous Plan-Excute-Analysis-Improve Cycle

Background-1
Modules in a QA System



Bacground-2: NTCIR-7 & -8 ACLIA



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

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

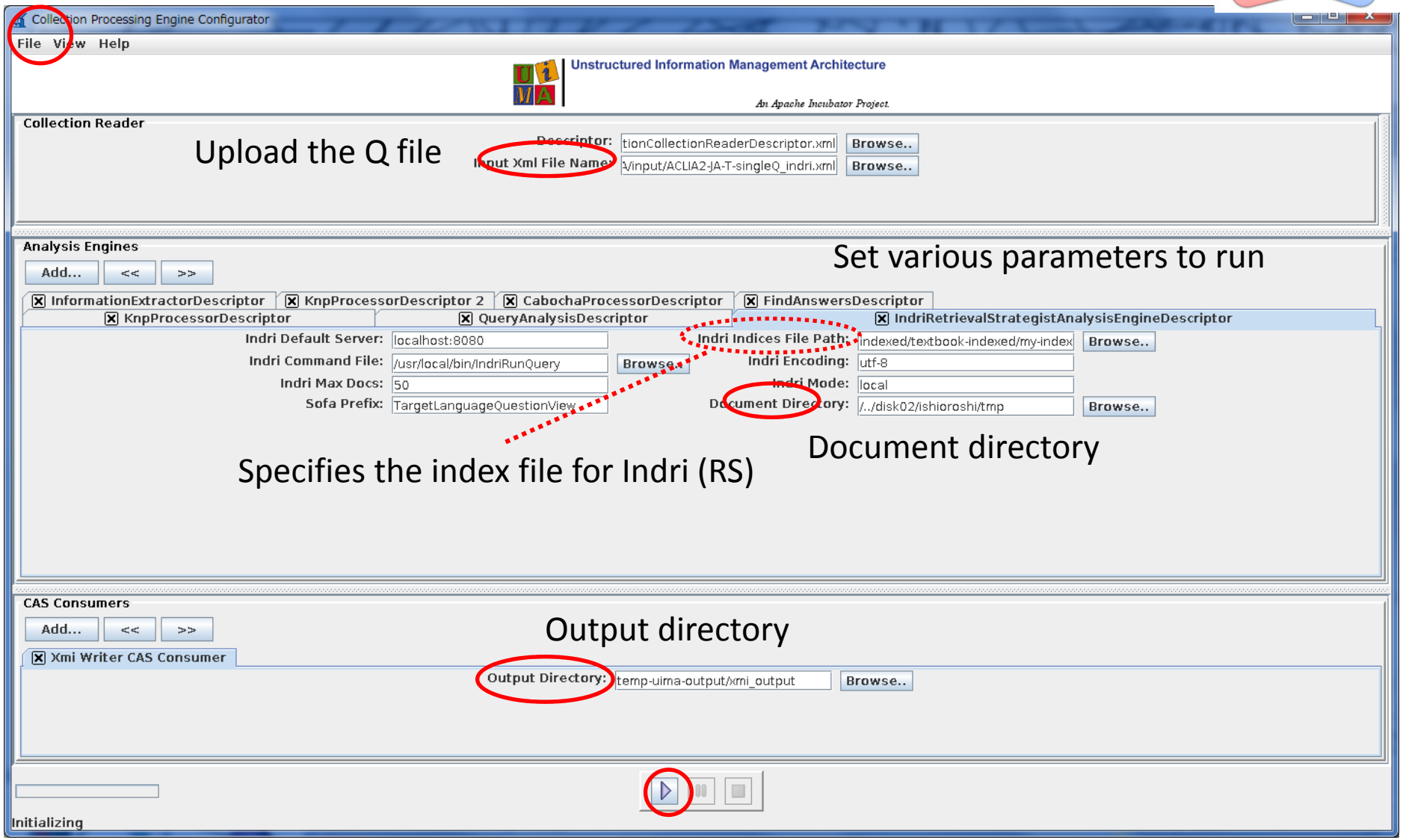
"Dream" System superior than Best End-to-End
but could not know why

http://aclia.lti.cs.cmu.edu/ntcir8/

Systems

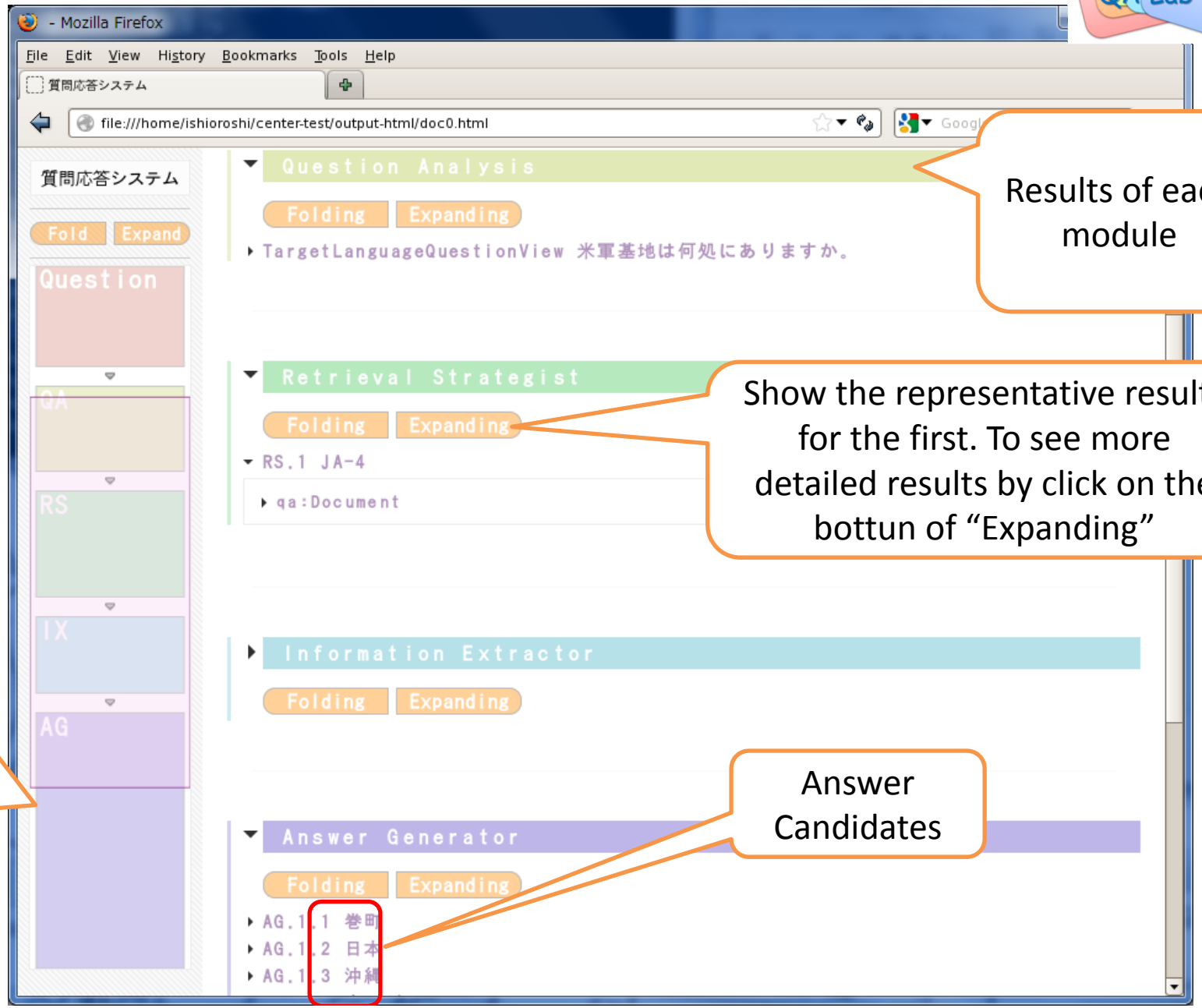
- QA-Platform
  - UIMA-based Module-base QA system
  - Follows: ACLIA modules
    - Q analysis
    - IR
    - Answer Extraction
    - Answer Generation
  - 2 baseline systems
    - Javelin (CMU)
    - MinerVA (YNU)
  - Systems are usable either
    - local (by Kachaco) or on NII Server “bunbun”

UIMA Native UI for QA Platform to run a pipeline



NTCIR

UI for the QA Platform to examine the results of each module



Corpus

- Questions & Answers
  - Japanese National Center for University Admissions Tests (multiple choices)
    - → use as a Yes-No Questions
    - XML, Japanese and English translation
  - Second Exam for the U of Tokyo
    - → factoid & Complex Questions
    - Will be XML tagged
- Knowledge sources provided
  - High school Textbooks, Wikipedia
  - Ontology of World History
  - Ontology-annotated Textbook
- Tools
  - Wide-range of Language annotation tools
  - RITE resource and tools

NTCIR

Evaluation Method-3

Tasks

	Multiple Choice (Center Exam)			2nd Exam	
	Japanese	English	English	Japanese	Japanese
NTCIR-11	World Hitory (YN)			Factoid	
NTCIR-12	Biology (YN)		Reading Comprehension	Factoid	Complex

Evaluate QA end-to-end with Every Possible combination of the modules submitted so far and base systems

NTCIR

Tentative Schedule:

Multiple Opportunities to Run

	Oct-13	Nov-13	Dec-13	Jan-14	Feb-14	Mar-13	Apr-14	May-14	Jun-13	Jul-14	Aug-14	Sep-14
1st Round	runs	return	results R-table									
2nd Round				runs	return	results R-table						
3rd Round							runs	return	results R-table			
4th Round										runs	return	results R-table

Submitted all rounds are not mandatory.

NTCIR

Participant's merits

- Can participate in one phase or multiple phases (flexible participation)
- Can evaluate with Baseline or with other systems and to improve your system for the next phase
- Can see the effectiveness of your modules in various pipelines created by other teams or combination of other teams.

NTCIR

Join Us!