



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



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

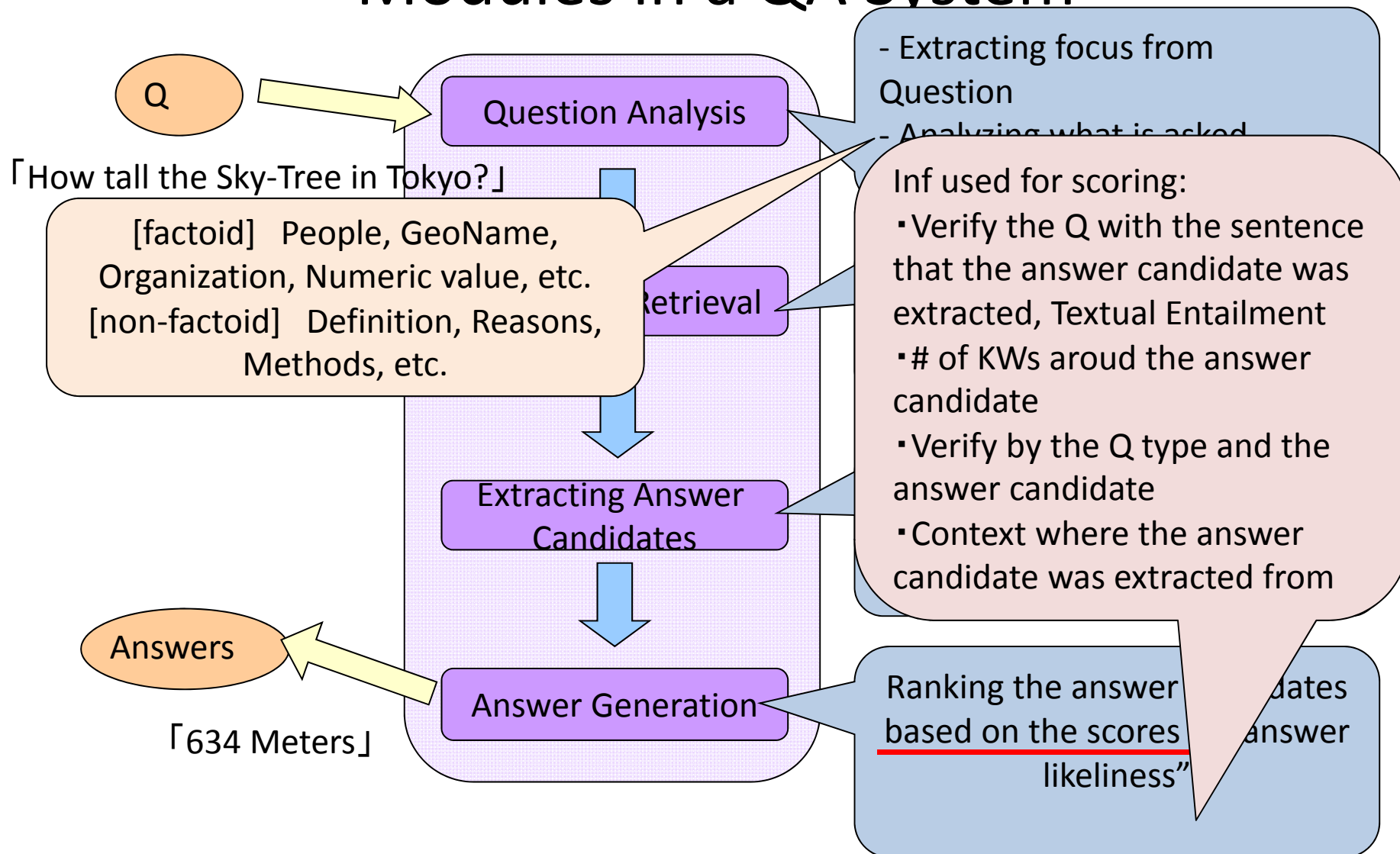


Main Idea

- 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

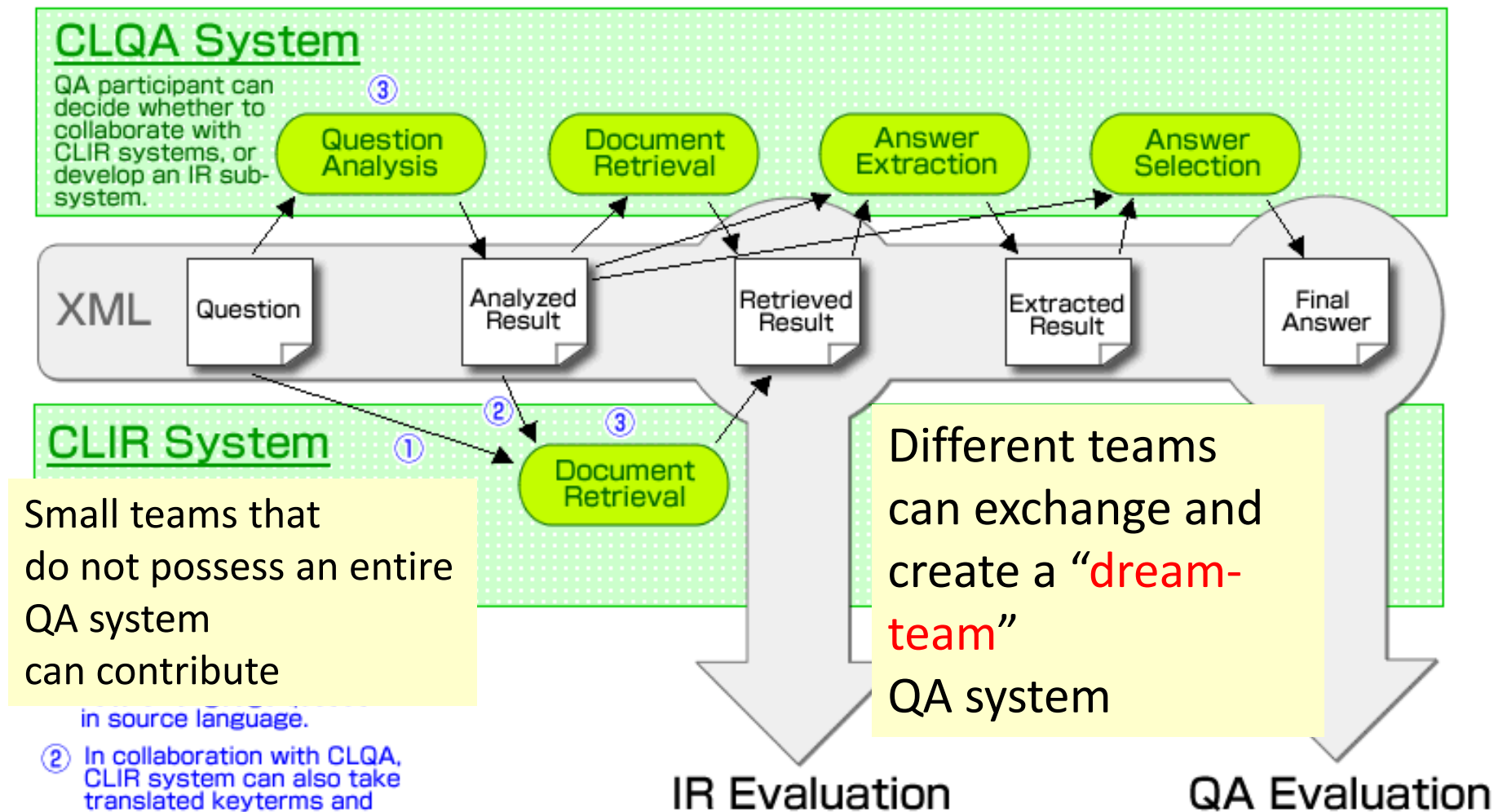
Background-1

Modules in a QA System





Sample System Architecture and Data Flow



<http://aclia.lti.cs.cmu.edu/ntcir8/>
NTCIR-8 overview 20100616

“Dream” System superior than Best End-to-End, but could not know why
Watariko Kando



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



Collection Processing Engine Configurator

File View Help

Unstructured Information Management Architecture
An Apache Incubator Project

Collection Reader

Upload the Q file

Descriptor:

Input Xml File Name:

Analysis Engines

Set various parameters to run

InformationExtractorDescriptor KnpProcessorDescriptor 2 CabochaProcessorDescriptor FindAnswersDescriptor

KnpProcessorDescriptor QueryAnalysisDescriptor IndriRetrievalStrategistAnalysisEngineDescriptor

Indri Default Server:

Indri Command File:

Indri Max Docs:

Sofa Prefix:

Indri Indices File Path:

Indri Encoding:

Indri Mode:

Document Directory:

Specifies the index file for Indri (RS)

Document directory

CAS Consumers

Output directory

Xmi Writer CAS Consumer

Output Directory:

Initializing

Run

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



The screenshot shows a Mozilla Firefox browser window displaying the QA Platform interface. The browser's address bar shows the file path: `file:///home/ishioroshi/center-test/output-html/doc0.html`. The page title is "質問応答システム". The interface is organized into a sidebar on the left and a main content area on the right. The sidebar contains a "Question" section and a vertical stack of colored buttons labeled "QA", "RS", "IX", and "AG". The main content area displays the results of four modules: "Question Analysis", "Retrieval Strategist", "Information Extractor", and "Answer Generator". Each module has a header bar with its name and two buttons: "Folding" and "Expanding". The "Retrieval Strategist" module is expanded, showing a sub-section "RS.1 JA-4" with a "qa:Document" entry. The "Answer Generator" module is also expanded, showing a list of answer candidates: "AG.1.1 巻頭", "AG.1.2 日本", and "AG.1.3 沖縄".

Results of each module

Show the representative results for the first. To see more detailed results by click on the bottun of "Expanding"

Click to jump to the module

Answer Candidates





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



Tasks

| | Multiple Choice (Center Exam) | | | 2nd Exam | |
|----------|-------------------------------|---------|-----------------------|----------|----------|
| | Japanese | English | English | Japanese | Japanese |
| NTCIR-11 | World History (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



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.



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.

Join Us!

