

NTCIR-12 QA-Lab Task second Pilot



Hideyuki Shibuki*1, Kotaro Sakamoto*1,*2,
Yoshionobu Kano*3,*4, †, Teruko Mitamura*5,
Madoka Ishioroshi*2, Tatsunori Mori*1, Noriko
Kando*2,*6

*1: Yokohama National University, *2: National Institute of
Informatics, *3: Shizuoka University, *4: PRESTO, *5: Carnegie
Mellon University, *6: The Graduate University for Advanced
Studies (SOKENDAI)



Challenges in QA Lab

- **Real-world** Question Answering using **university entrance exam** on World History
 - Questions are often multiple sentences
 - Context
 - Various Question-Format (multiple choices, Yes/No, Slot Filling, Chronological Ordering, Essay, Combo)
 - Distribution of the Question Types is unknown and vary year by year
 - Some questions require inference
- Explore the role of **Semantic** Processing
- A **module-based platform** for advanced question answering systems
- Multiple Phases (you can try more than one time)
- Japanese and English

Question XML Format

第1問 人類が営む生業と労働は、経済・社会・政治の動きと密接にかかわりながら、大きく変容してきた。生業と労働の歴史について述べた次の文章A～Cを読み、下の問い(問1～9)に答えよ。(配点 25)

A 清の学者趙翼は、明代の文化人の趨勢を論じて、①唐宋以来、文化・芸術に秀でた者の多くは科擧の合格者であったが、②明代になってその担い手は在野の人物に移っていったと述べている。明代中期の画家唐寅は、まさにその過渡期の人物と言える。彼は科擧で優秀な成績を収めながらも、不運な事件に巻き込まれ、栄達の道を絶たれてからは、蘇州で画業をなりわいとしながら自由奔放な生活を送った。明代中期から後期にかけて、在野の芸術家や文筆家が続々と現れたのは、③江南を中心とする商工業の発展によって都市の文化が成熟し、絵画や出版物が広く商品としての価値を持つようになったからであった。

問1 下線部①に関連して、次に挙げる人物は、いずれも唐代から宋代にかけての科擧の合格者である。それぞれの人物について述べた文として正しいものを、次の①～④のうちから一つ選べ。

- ① 歐陽脩や蘇軾は、唐代を代表する文筆家である。
- ② 顔真卿は、宋代を代表する書家である。
- ③ 宋の王安石は、新法と呼ばれる改革を行った。
- ④ 秦檜は、元との関係をめぐり主戦派と対立した。

```

<exam source="National Center For University Entrance Examination" subject="SekaishiB(main exam)"
year="2009">
Center-2009--Main-SekaishiB<br/>
<title>
2009年度 本試験 世界史B<br/><br/>
</title>
<question id="Q1" minimal="no">
<label>【1】</label>
<instruction>
<br/>人類が営む生業と労働は、経済・社会・政治の動きと密接にかかわりながら、大きく変容してき
た。生業と労働の歴史について述べた次の文章A～Cを読み、以下の問い(問1～9)に答えよ。<br/> (配
点 25)<br/>
</instruction>
<data id="D0" type="text">
<label>A</label><br/> 清の学者趙翼は、明代の文化人の趨勢を論じて、<uText
id="U1"><label>(1)</label>唐宋以来、文化・芸術に秀でた者の多くは科擧の合格者であった</uText>が、
<uText id="U2"><label>(2)</label>明代になってその担い手は在野の人物に移っていったと述べ
ている。明代中期の画家唐寅は、まさにその過渡期の人物と言える。彼は科擧で優秀な成績を収めなが
らも、不運な事件に巻き込まれ、栄達の道を絶たれてからは、蘇州で画業をなりわいとしながら自由奔放
な生活を送った。明代中期から後期にかけて、在野の芸術家や文筆家が続々と現れたのは、<uText
id="U3"><label>(3)</label>江南を中心とする商工業の発展</uText>によって都市の文化が成熟し、絵画
や出版物が広く商品としての価値を持つようになったからであった。<br/><br/>
</data>
<question anscol="A1" answer_style="multipleChoice" answer_type="sentence" id="Q2"
knowledge_type="KS" minimal="yes">
<label>問1</label>
<instruction>
下線部<ref comment="" target="U1">(1)</ref>に関連して、次に挙げる人物は、いずれも唐代から宋代に
かけての科擧の合格者である。それぞれの人物について述べた文として正しいものを、次の①～④のう
ちから一つ選べ。
</instruction>
<ansColumn id="A1">1</ansColumn><br/>
<choices anscol="A1" comment="">
<choice ansnum="1">
<cNum>①</cNum> 歐陽脩や蘇軾は、唐代を代表する文筆家である。</choice>
<choice ansnum="2">
<cNum>②</cNum> 顔真卿は、宋代を代表する書家である。</choice>
<choice ansnum="3">
<cNum>③</cNum> 宋の王安石は、新法と呼ばれる改革を行った。</choice>
<choice ansnum="4">
<cNum>④</cNum> 秦檜は、元との関係をめぐり主戦派と対立した。</choice><br/></choices>
</question>
.....
</exam>

```

Question XML Format

第1問 人類が営む生業と労働は、経済・社会・政治の動きと密接にかかわりながら、大きく変容してきた。生業と労働の歴史について述べた次の文章A～Cを読み、下の問い(問1～9)に答えよ。(配点 25)

A 清の学者趙翼は、明代の文化人の趨勢を論じて、①唐宋以来、文化・芸術に秀でた者の多くは科擧の合格者であったが、②明代になってその担い手は在野の人物に移っていったと述べている。明代中期の画家唐寅は、まさにその過渡期の人物と言える。彼は科擧で優秀な成績を収めながらも、不運な事件に巻き込まれ、栄達の道を絶たれてからは、蘇州で画業をなりわいとしながら自由奔放な生活を送った。明代中期から後期にかけて、在野の芸術家や文筆家が続々と現れたのは、③江南を中心とする商工業の発展によって都市の文化が成熟し、絵画や出版物が広く商品としての価値を持つようになったからであった。

問1 下線部①に関連して、次に挙げる人物は、いずれも唐代から宋代にかけての科擧の合格者である。それぞれの人物について述べた文として正しいものを、次の①～④のうちから一つ選べ。

- ① 歐陽脩や蘇軾は、唐代を代表する文筆家である。
- ② 顔真卿は、宋代を代表する書家である。
- ③ 宋の王安石は、新法と呼ばれる改革を行った。
- ④ 秦檜は、元との関係をめぐり主戦派と対立した。

2015-02-27

qalab rtm

Questions (Multi-Sentence, Suggest Context)

```

<exam source="National Center For University Entrance Examination" subject="SekaishiB(main exam)"
year="2009">
Center-2009--Main-SekaishiB<br/>
<title>
2009年度 本試験 世界史B<br/><br/>
</title>
<question id="Q1" minimal="no">
<label>【1】</label>
<instruction>
<br/><br/>
人類が営む生業と労働は、経済・社会・政治の動きと密接にかかわりながら、大きく変容して
きた。生業と労働の歴史について述べた次の文章A～Cを読み、以下の問い(問1～9)に答えよ。
(配点 25)<br/>
</instruction>
<data id="D0" type="text">
<label>A</label><br/>
清の学者趙翼は、明代の文化人の趨勢を論じて、<uText
id="U1"><label>(1)</label>唐宋以来、文化・芸術に秀でた者の多くは科擧の合格者であった</uText>が、
<uText id="U2"><label>(2)</label>明代</uText>になってその担い手は在野の人物に移っていったと述
べている。明代中期の画家唐寅は、まさにその過渡期の人物と言える。彼は科擧で優秀な成績を収めな
がらも、不運な事件に巻き込まれ、栄達の道を絶たれてからは、蘇州で画業をなりわいとしながら自由放
散な生活を送った。明代中期から後期にかけて、在野の芸術家や文筆家が続々と現れたのは、<uText
id="U3"><label>(3)</label>江南を中心とする商工業の発展</uText>によって都市の文化が成熟し、絵画
や出版物が広く商品としての価値を持つようになったからであった。<br/><br/>
</data>
<question anscol="A1" answer_style="multipleChoice" answer_type="sentence" id="Q2"
knowledge_type="KS" minimal="yes">
<label>問1</label>
<instruction>
下線部<ref comment="" target="U1">(1)</ref>に関連して、次に挙げる人物は、いずれも唐代から宋代に
かけての科擧の合格者である。それぞれの人物について述べた文として正しいものを、次の①～④のう
ちから一つ選べ。
</instruction>
<ansColumn id="A1">1</ansColumn><br/>
<choices anscol="A1" comment="">
<choice ansnum="1">
<cNum>①</cNum> 歐陽脩や蘇軾は、唐代を代表する文筆家である。</choice>
<choice ansnum="2">
<cNum>②</cNum> 顔真卿は、宋代を代表する書家である。</choice>
<choice ansnum="3">
<cNum>③</cNum> 宋の王安石は、新法と呼ばれる改革を行った。</choice>
<choice ansnum="4">
<cNum>④</cNum> 秦檜は、元との関係をめぐり主戦派と対立した。</choice><br/></choices>
</question>
.....
</exam>

```

4

Question XML Format 1 (Center Exam)



第1問 人類が営む生業と労働は、経済・社会・政治の動きと密接にかかわりながら、大きく変容してきた。生業と労働の歴史について述べた次の文章A～Cを読み、下の問い(問1～9)に答えよ。(配点 25)

A 清の学者趙翼は、明代の文化人の趨勢を論じて、①唐宋以来、文化・芸術に秀でた者の多くは科擧の合格者であったが、②明代になってその担い手は在野の人物に移っていったと述べている。明代中期の画家唐寅は、まさにその過渡期の人物と言える。彼は科擧で優秀な成績を収めながらも、不運な事件に巻き込まれ、栄達の道を絶たれてからは、蘇州で画業をなりわいとしながら自由奔放な生活を送った。明代中期から後期にかけて、在野の芸術家や文筆家が続々と現れたのは、③江南を中心とする商工業の発展によって都市の文化が成熟し、絵画や出版物が広く商品としての価値を持つようになったからであった。

問1 下線部①に関連して、次に挙げる人物は、いずれも唐代から宋代にかけての科擧の合格者である。それぞれの人物について述べた文として正しいものを、次の①～④のうちから一つ選べ。

- ① 歐陽脩や蘇軾は、唐代を代表する文筆家である。
- ② 顔真卿は、宋代を代表する書家である。
- ③ 宋の王安石は、新法と呼ばれる改革を行った。
- ④ 秦檜は、元との関係をめぐり主戦派と対立した。

Questions (Multi-Sentence, Suggest Context)

```
year="2009">
Center-2009--Main-SekaishiB<br/>
<title>
2009年度 本試験 世界史B<br/><br/>
</title>
```

Context

```
点 25<br/>
</instruction>
<data id="D0" type="text">
<label>A</label><br/> 清の学者趙翼は、明代の文化人の趨勢を論じて、<uText
id="U1"><label>(1)</label>唐宋以来、文化・芸術に秀でた者の多くは科擧の合格者であった</uText>が、
<uText id="U2"><label>(2)</label>明代</uText>になってその担い手は在野の人物に移っていったと述べて
いる。明代中期の画家唐寅は、まさにその過渡期の人物と言える。彼は科擧で優秀な成績を収めなが
らも、不運な事件に巻き込まれ、栄達の道を絶たれてからは、蘇州で画業をなりわいとしながら自由奔放
```

Sub-Questions

```
<question anscol="A1" answer_style="multipleChoice" answer_type="sentence" id="Q2"
knowledge_type="KS" minimal="yes">
<label>問1</label>
<instruction>
下線部<ref comment="" target="U1">(1)</ref>に関連して、次に挙げる人物は、いずれも唐代から宋代に
かけての科擧の合格者である。それぞれの人物について述べた文として正しいものを、次の①～④のう
ちから一つ選べ。
</instruction>
<ansColumn id="A1">1</ansColumn><br/>
```

Multiple Choices

```
<choice ansnum="3">
<cNum>③</cNum> 宋の王安石は、新法と呼ばれる改革を行った。</choice>
<choice ansnum="4">
<cNum>④</cNum> 秦檜は、元との関係をめぐり主戦派と対立した。</choice><br/></choices>
</question>
.....
</exam>
```



What's New in NTCIR-12

- **Increasing Training Data**
 - 1 year (NTCIR-11)
 - → 3 years (NTCIR-12) + some additions for specific question type
 - → Joint Effort to Construct Large Scale Quasi-Examples
 - By Extracting sentences from Textbook, Collecting data from Exercise Books or related test, creating by Human experts, etc.
- Training and 1 test by **Question Types** (New)
- 1 Test use a whole questions in an examination (same as NTCIR-11)
- Vertical Combinations of Different modules or QA Systems (and explore more complex ways for better integration)



Question Types Proposed

- 【A1:大論述 (Complex Essay, CE)】
- 【A2:小論述 (Simple Essay, SE)】
- 【B1:Factoid (F)】
- 【B2:穴埋め (Slot-Filing, SF)】
- 【C:真偽判定型 (True-or-False, TF)】
- 【D:特殊 (Unique, U)】



Data and Resources

- Questions and Answers
 - Training data – Three Years of Examinations (National Center Exam, Secondary Exam for 5 Univ)
 - Test data – 2 Sets
- Knowledge Sources
 - High School Textbook (4 types from 2 publishers, Japanese Only)
 - Wikipedia
 - Participants can use any resources (and need to report)



Tools Provided

- Two Baseline Systems (Japanese and English)
 - UIMA module-based end-to-end QA systems
- One Passage Retrieval Module
 - to enhance the module-based collaboration.



Evaluation

- Non-Essay
 - Scores set by National Center or Each University
 - Accuracy
- Essay
 - Rouge
 - BE-based automatic quasi-pyramid
 - Pyramid using human extracted nuggets and nugget voting (30 Complex and 50 Simple Essays)
- **Welcome! More discussion about the evaluation methodologies and metrics**



Proposed Schedule -1

- April – June 2015
Discussion on task description, Question types ontology, Evaluation Methodologies and metrics
- End June 2015 Fixed task detailed task description



Proposed Schedule -2

<Phase 1>

- July 2015 1st runs
- Aug 2015 Results of the 1st runs return to the participants
- Sept2015 Round-table discussion on the above results

<Phase 2>

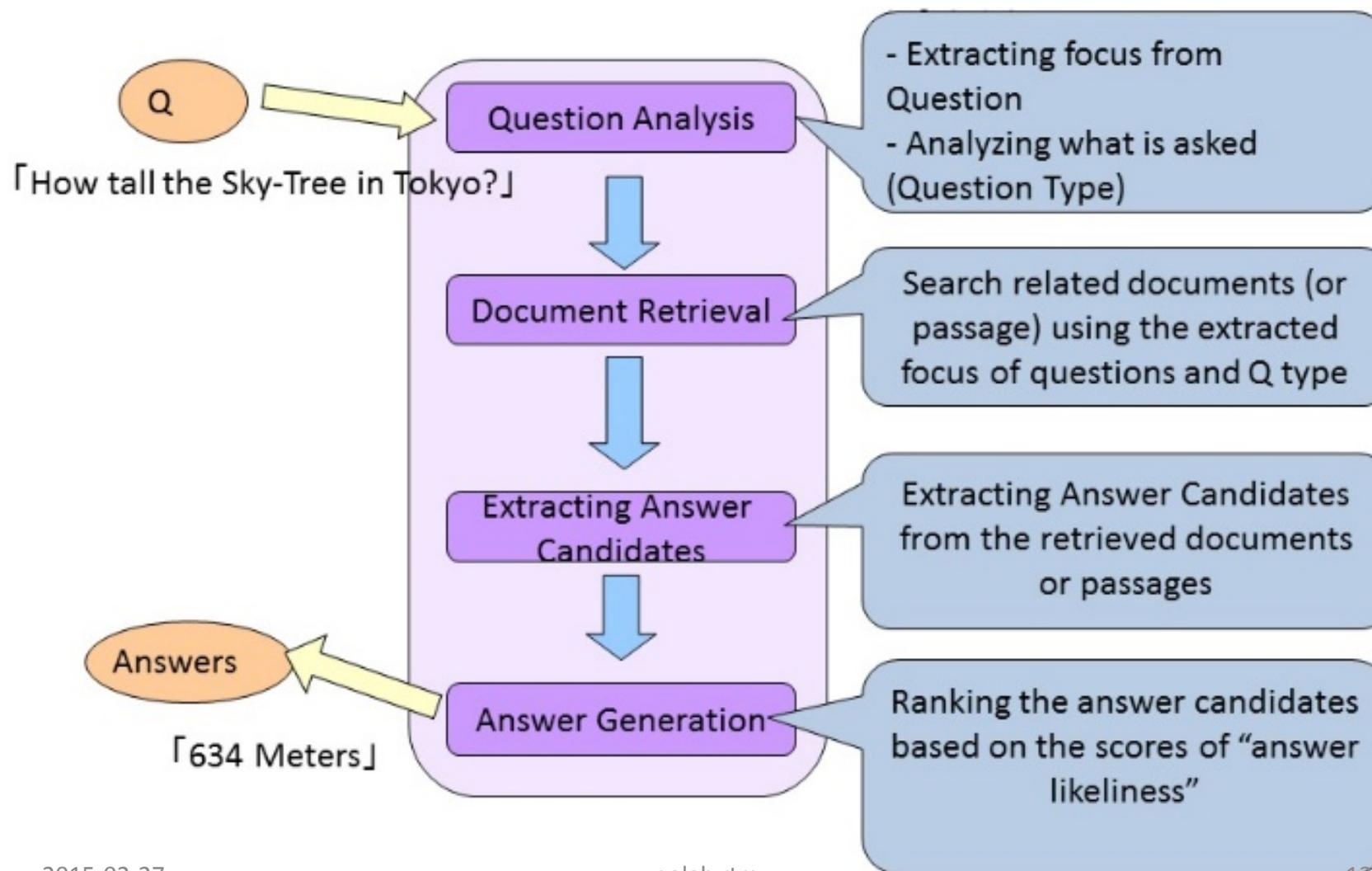
- Dec 2015 2nd runs
- Jan 2016 Results of the 2nd runs return to the participants
- Feb 2016 Round-table discussion on the above results

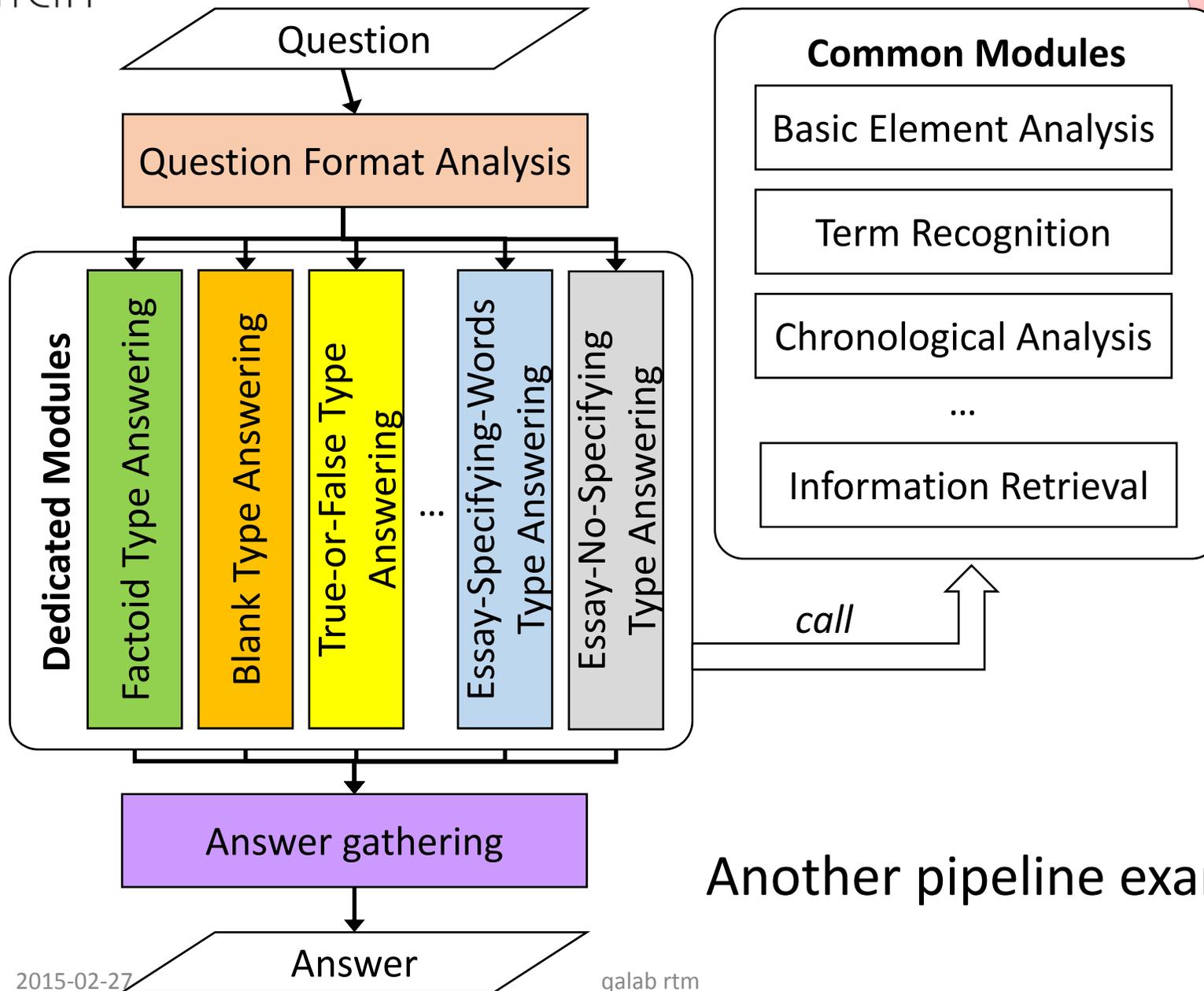
<Towards Conference>

- May 1: 2016 Final Paper submission
- June 2016 NTCIR-12 Conference

Module Structure

(Pipeline planned in NTCIR-11)





Another pipeline example



Join Us !!

Email: qalab-admin@nii.ac.jp