

# SML Question-Answering System for World History Essay and Multiple-choice Exams at NTCIR-12 QA Lab-2

**Team:SML**

**Takuma Takada, Takuya Imagawa, Takuya Matsuzaki, Satoshi Sato  
Graduate School of Engineering, Nagoya University**

# Our Challenge

- National Center Test
  - Multiple-choice type questions
- Second-stage Examinations ← today's presentation
  - Short-essay type questions
  - Long-essay type questions
  - Factoid type questions

# World History Short Essay

Example	Reference
キリスト教はローマ皇帝によってどのように公認されたか、その皇帝の名前と公認の理由(reason)に触れながら、60字以内で説明しなさい。	Tokyo University 2013-2-(1)b
明から清の前期にかけて、対外貿易と朝貢との関係がどのように変化(change)したかについて、海禁政策に着目しながら、120字以内で説明しなさい。	Tokyo University 2011-2-(2)
ギリシア人は文化的には一つの民族であるという共通の認識を支えた諸要素を、60字以内で説明しなさい。	Tokyo University 2009-2-(1)b
インド亜大陸へのイスラームの定着は海陸両方の経路から進化した。そのうち、カイバル峠を透ルートによる定着過程の10世紀末から16世紀前半にかけての展開(process)を、政治的側面と文化的側面の双方にふれながら120字以内で説明しなさい。	Tokyo University 2006-2-(1)
ポリスの形成過程(process)を、60字以内で説明しなさい。	Tokyo University 2009-2-(1)b

# World History Short Essay

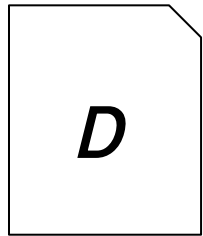
Example	Reference
<p>キリスト教はローマ皇帝によってどのように公認されたか、その皇帝の名前と公認の理由(reason)に触れながら、<b>60字以内</b>で説明しなさい。</p>	Tokyo University 2013-2-(1)b
<p>明から清の前期にかけて、対外貿易と朝貢との関係がどのように変化(change)したかについて、海禁政策に着目しながら、<b>120字以内</b>で説明しなさい。</p>	Tokyo University 2011-2-(2)
<p>ギリシア人は文化的には一つの民族であるという共通の認識を支えた諸要素を、<b>60字以内</b>で説明しなさい。</p>	Tokyo University 2009-2-(1)b
<p>インド亜大陸へのイスラームの定着は海陸両方の経路から進化した。そのうち、カイバル峠を透ルートによる定着過程の10世紀末から16世紀前半にかけての展開(deployment)を、政治的側面と文化的側面の双方にふれながら<b>120字以内</b>で説明しなさい。</p>	Tokyo University 2006-2-(1)
<p>ポリスの形成過程(process)を、<b>60字以内</b>で説明しなさい。</p>	Tokyo University 2009-2-(1)b

# Approach

- We developed a system that
  - Generates an answer by extracting sentences
  - Searches based on the surface similarity
  - Compresses to fit within the character limit

## Question

キリスト教はローマ皇帝によってどのように公認されたか、その皇帝の名前と公認の理由に触れながら、**60字以内**で説明しなさい。



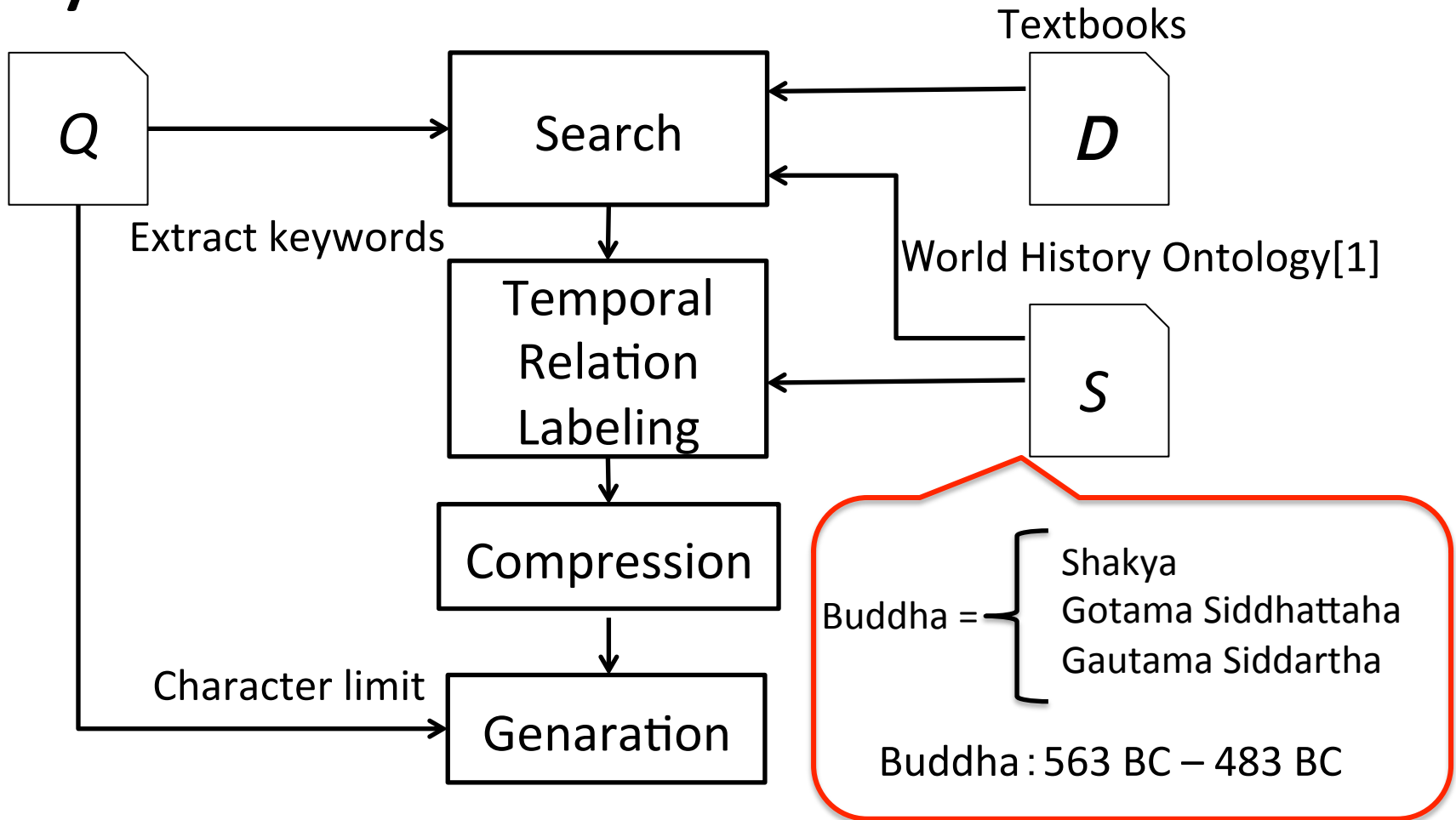
Search →

## Output

迫害よりも懐柔が得策と考えるにいたった  
コンスタンティヌス帝は313年ミラノ勅令によって  
キリスト教を公認した。

Textbooks ~ 20,000 sentences

# System



[1]Kawazoe, A. et.al. (2014). “World History Ontology for Reasoning Truth/Falsehood of Sentences: Event Classification to Fill in the Gaps between Knowledge Resources and Natural Language Texts,” Lecture Notes in Computer Science 8417, pp.42-50

# Seaching based on Surface Similarity[2]

Question

Tokyo university 2013-2-(1)-(b)

キリスト教はローマ皇帝によってどのように公認されたか、その皇帝の名前と公認の理由に触れながら、60字で説明しなさい。



Extract nouns

**Match?**

Sentence in textbooks



Extract nouns

迫害よりも懐柔が得策と考えるにいたったコンスタンティヌス帝は313年ミラノ勅令によってキリスト教を公認した。

[2] Shohei Hattori and Satoshi Sato.

“Surface-similarity based textual entailment recognition for japanese text.”

JSAI Journal,, 2014, vol.29-4, pp416-426

# Seaching based on Surface Similarity[2]

Question

Tokyo university 2013-2-(1)-(b)

キリスト教はローマ皇帝によってどのように公認されたか、その皇帝の名前と公認の理由に触れながら、60字で説明しなさい。



Extract nouns

**Match?**

Sentence in textbooks



Extract nouns

迫害よりも懐柔が得策と考えるにいたったコンスタンティヌス帝は313年ミラノ勅令によってキリスト教を公認した。

[2] Shohei Hattori and Satoshi Sato.

“Surface-similarity based textual entailment recognition for japanese text.”

JSAI Journal,, 2014, vol.29-4, pp416-426



# Seaching based on Surface Similarity[2]

$$score = \frac{\sum_{x \in \mathbf{N}^1} \{\min(f(x, t), f(x, q)) \cdot a(x) + 0.3 \cdot c\}}{\sum_{x \in \mathbf{N}^1} f(x, q)}$$

t : textbook  
q: question

$$a(x) = \begin{cases} 1 & \text{if } x \in \text{Headword of the glossary} \\ 0.7 & \text{else if } x \in \text{Named entities} \\ 0.5 & \text{else if } x \in \text{Wikipedia page title} \\ 0.1 & \text{otherwise.} \end{cases}$$

[2] Shohei Hattori and Satoshi Sato.

“Surface-similarity based textual entailment recognition for japanese text.”

JSAI Journal,, 2014, vol.29-4, pp416-426

# Temporal Relation Labeling

## Question

5世紀におけるフン族の最盛期とその後について、60字以内で説明しなさい。  
(5<sup>th</sup> century) (Huns)

## Sentence in textbooks

最盛期をきずいたアッティラの死後王国は分裂を重ねて弱体化しフン族は  
(Attila) (Huns)

アヴァール人やブルガール族などに吸収されていった。  
(Avars) (Bulgars)

## temporal information in the question

5<sup>th</sup> century 401 - 500

Huns 301 - 701

## temporal information in a sentence

Attila 406 - 453

Huns 301 - 701

⋮

- Classify textbook sentences into three types, Match, Unknown, or Mismatch



**Compare at least  
one common  
temporal information**

# Necessity of Compression

Question

Keio 2015 I 3(3)

1884年から1885年にかけて開催されたベルリン会議を契機に、奴隷の供給地であったアフリカの植民地化が加速された。この会議で定められたアフリカの分割・植民地化に関する原則について、**30字**以内で説明しなさい。

Output

a sentence in textbooks average 54 characters

ベルギーのコンゴ支配をめぐってドイツのビスマルクの提唱で、1884年に開かれたベルリン会議では、アフリカ大陸を分割する上での原則(無主地の有効な占有)が決定され、侵略競争が公然と始まる。

⋮

**99 characters** score=**1.18**

宗主国の清はこれを認めず清仏戦争が起こった。

**29 characters** score=**0.13**

# Compression

When the score < threshold

→ recalculate the score for each segment  
separated by commas

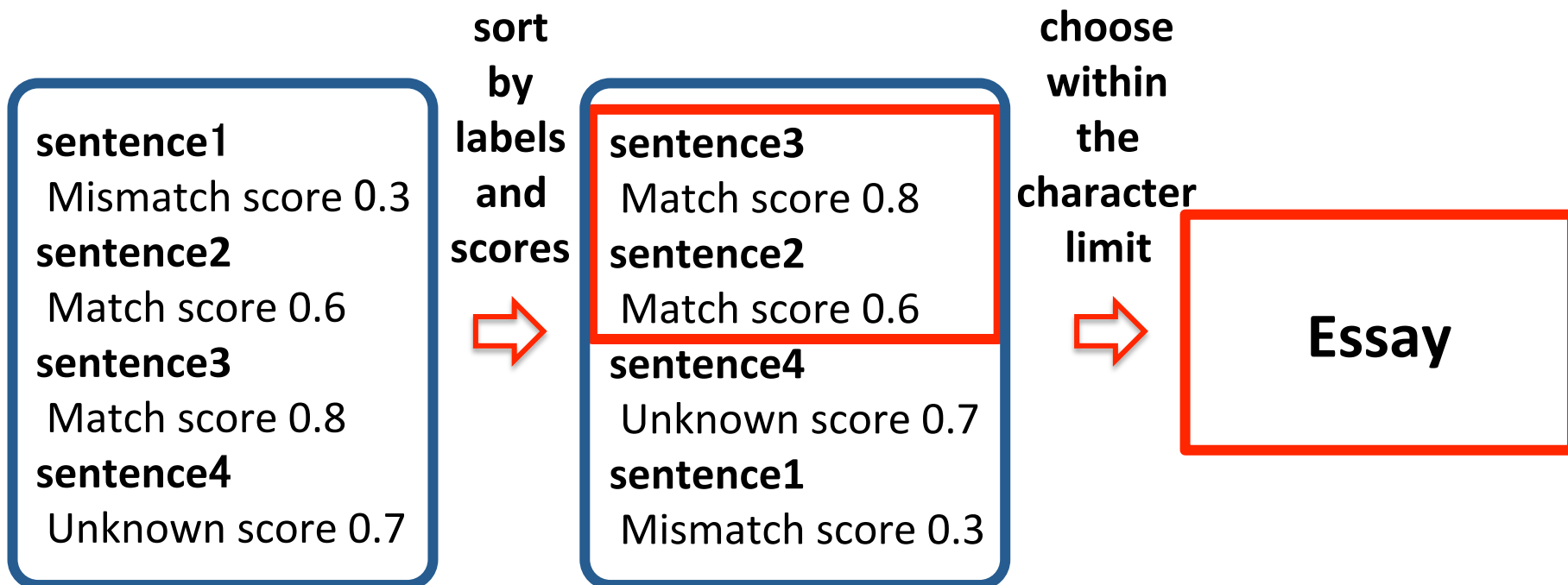
ベルギーのコンゴ支配をめぐってドイツのビスマルクの提唱で、1884年に開かれたベルリン会議では、**アフリカ大陸を分割する上での原則(無主地の有効な占有)が決定され、**侵略競争が公然と始まる。

Output

calculate the score again

**アフリカ大陸を分割する上での原則(無主地の有効な占有)が決定**

# Generation



# Long essay

# World History Long Essay

Example	Reference
<p>7世紀から18世紀にかけての西ヨーロッパ・西アジア・東アジアでは、整備された統治機構と強力な軍隊を持ち、それぞれ異なる思想・理念に基づく集権的な体制を持つ国家が確立し、各地域の政治・外交面で中心的役割を担った。フランスやこれと外交面で提携したオスマン帝国、清朝がその典型である。しかし、18世紀に入るとそれぞれの地域における社会の変動から集権的支配に緩みが生じるとともに、あるいは国家体制を支える理念に対抗する思想も生まれた。そして、18世紀末には国家体制を動揺させる事件も起こった。以上の点を踏まえ、17世紀から18世紀末までの西ヨーロッパ・西アジア・東アジアにおける集権的国家体制の変遷について、地域ごとの思想・理念および、体制維持に影響を及ぼした宗教の役割にも留意しつつ、略述しなさい。解答は、600文字以内で記述し、必ず次の9つの語句を一度は用いて、その語句に下線を付しなさい。</p> <p>指定語句：ウェストファリア条約, ウラマー, クリム=ハン国, 啓蒙思想, チベット仏教, ナントの王令廃止, 白蓮教徒の乱, 満漢併用制, メッカ</p>	<p>Phase2 Sundai Exam 2015-Aug I</p>

# Factoid



# Factoid-type Questions

Question

Sundai Exam 問1

神聖文字を解読したフランスのエジプト学者の名前を記しなさい。

↓  
Category  
Prediction

Glossary categorized

Searching based on surface similarity

D

シャンポリオン  
1790年から1832年  
フランスのエジプト学者。  
1822年、ロゼッタ=ストーンの  
碑文から神聖文字の  
解読に成功した。

Extraction entry name

Output

シャンポリオン

# Category Prediction

唐の僧・義浄がインドから中国に帰る途中、この海峡の南に位置するスマトラ島のシャリーヴィジャヤ王国に滞在して書き記した旅行記の名称を記しなさい。



## 義浄

唐僧。671年広州を出発し、南海経由でインドへ行き、ナーランダ一僧院で学び、695年に帰国した。

旅行記「南海寄帰内法伝」を帰国途中のシャリーヴィジャヤで著した。

## 南海寄帰内法伝

義浄のインド・南海の旅行記。

4巻。インドからの帰途滞在したシャリーヴィジャヤで記述され、インド史・南海史・仏教史研究の重要史料とされる。

# Category Prediction

唐の僧・義浄がインドから中国に帰る途中、この海峡の南に位置するスマトラ島のシャリーヴィジャヤ王国に滞在して書き記した旅行記の名称を記しなさい。

Person



Book



Book



義浄

唐僧。671年広州を出発し、南海経由でインドへ行き、ナーランダ一僧院で学び、695年に帰国した。

旅行記「南海寄帰内法伝」を帰国途中のシャリーヴィジャヤで著した。

南海寄帰内法伝

義浄のインド・南海の旅行記。4巻。インドからの帰途滞在したシャリーヴィジャヤで記述され、インド史・南海史・仏教史研究の重要史料とされる。

# Category Prediction

- Classifies the question into the 21 types

神聖文字を解読したフランスのエジプト**学者**の名前を記しなさい。

(Scholar)



category : **Person**

Dictionary-based: 1036 noun-category associations  
extracted from wikipedia pages for world-history terms.

evt:人物(h)	エラスムス	1469頃
evt:人物(h)	エラトステネス	前275
evt:人物(h)		1533

(Eratosthenes)

家・王朝 人物

(Person)

world history ontology

エラトステネス (Eratosthenes)

(Ερατοσθένης, Eratosthenes, 紀元前275年 - 紀元前194年) は、ヘレニズム時代のエジプトで活躍したギリシャ人の**学者**であり、アレクサンドリア図書館を含む研究**(Scholar)**ムセイオンの**館長**を務めた。業績は文献学、地理学を始めヘレニ

(Director)

wikipedia page

# Result of Sundai mock test

Phase1	No.	Point
Long essay	1	4/28
Short essay	1a	0/6
	1b	1/5
	2a	0/3
	2b	0/4
	3	0/4
Factoid	1-10	2/10
total		7/60

Phase2	No.	Point
Long essay	1	3/26
Short essay	1a	3/6
	1b	0/4
	2	1/4
	3a	0/4
	3b	0/4
Factoid	1-10	8/10
total		15/60

# Issue : mismatch of nouns by the difference of abstraction levels

## Question

Sundai Exam 3(a)

宋代における都市の発達やその商業活動について  
唐代と対比しながら、60字以内で説明しなさい。

## Output of our system

Point 0/4

それまでの重要な交易相手であったローマ帝国やササン朝が  
崩壊したため、商業活動が沈滞し、都市も衰退した。

## the most appropriate sentence in textbooks

商品流通はますます大規模になり、行(商人)・作(手工業者)など  
の同業組合もうまれ、塩・茶・米・絹などを扱う大商人が活躍した。

Question (abstract description)

textbooks (concrete description)



# Issue: necessity of combination of several relevant sentences

## Question

アナトリア半島(小アジア)の西岸に位置するミレトスは, ...  
金属貨幣とギリシアの民主政治との関わりについて90字で説明しなさい。

## Output of our system

Point 1/6

人口も増加し土地が不足したので, 前8世紀半ばからギリシア人は大規模な植民活動に乗り出し, 地中海と黒海の沿岸各地に植民市を建設したが, これは貿易活動を活発化させることになった。



get 1 score by this description only

## relevant sentences

- リディアではじめて貨幣が鑄造され, ギリシア人もそれにならうことになった。
- ポリスの戦士は武具を自費で用意するのが原則であったが, 金属の輸入で武器が安くなると, 富裕な平民は武具を買って参戦できるようになった。
- これにより彼らが多数参加して密集隊形(ファランクス)を組んで戦う重装歩兵部隊が, 騎馬を利用する貴族にかわって軍隊の主力となった。

# Issue: necessity of combination of several relevant sentences

- 人口も増加し土地が不足したので、前8世紀半ばからギリシア人は大規模な植民活動に乗り出し、地中海と黒海の沿岸各地に植民市を建設したが、これは貿易活動を活発化させることになった。
- リディアではじめて貨幣が鑄造され、ギリシア人もそれにならうことになった。
- ポリスの戦士は武具を自費で用意するのが原則であったが、金属の輸入で武器が安くなると、富裕な平民は武具を買って参戦できるようになった。
- これにより彼らが多数参加して密集隊形(ファランクス)を組んで戦う重装歩兵部隊が、騎馬を利用する貴族にかわって軍隊の主力となった。



**Extract only  
the required parts (blue parts)**

**better essay**



# Summary

- Developed a world history Second-stage exam automatic answering system.
  - Answer generation by extractive summary using surface similarity.
  - Introduced restriction by temporal relation and compression.
  - Made synonyms and time information dictionaries.
- Issues in short-essay answering system
  - the mismatch of nouns by the difference of abstraction levels.
  - necessity of combination of several relevant sentences.