

WiseQA: 하이브리드 추론 기반 자연어 질의응답 기술



2015. 8. 21.

김현기/장명길/류법모/이형직/허정/배용진

지식마이닝연구실

한국전자통신연구원

발표 순서

I WiseQA 개요

II QA 관련연구

III WiseQA 연구 성과 및 계획

IV 맺음말

기술동향 및 연구 방법론 (요약)

기술 동향

- **IBM Watson:** 비정형 지식베이스 기반 귀납적 문제해결
 - 질문: 여러 개의 문장으로 구성된 **복잡한 질문**
 - 정답: **비정형 지식에서 62%, 정형 지식에서 7.8%** 찾음
- **Google Knowledge Graph:** 정형 지식베이스 기반 연역적 문제해결
 - 질문: 한 개 문장, 제한된 문형의 **단순한 질문**
 - 정답: 질문의 객체와 매칭되는 지식그래프의 **단순사실 정보 제공**

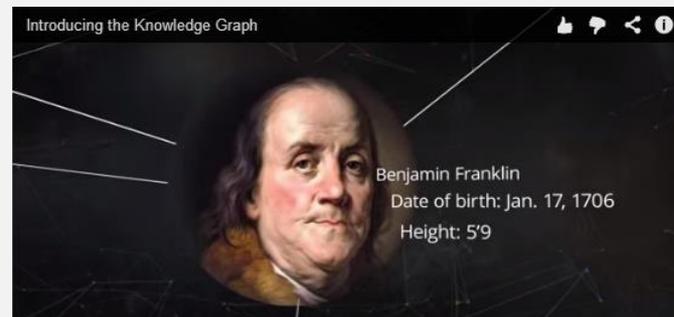
<IBM Watson (1천만권 도서 분량)>

	Document search	Passage search	Answer Lookup	PRISMATIC search	All
Percentage of questions in active subset	99.07%	100.00%	13.64%	43.75%	100%
No. of answers/active question	90.55	162.47	15.11	11.57	216.53
Binary recall	74.43%	79.40%	3.53%	8.31%	87.17%
Percentage unique	7.24%	12.05%	0.03%	0.18%	—
Accuracy	62.65%	62.74%	2.18%	5.65%	71.32%

비정형 지식베이스: 62%

정형 지식베이스: 7.8%

<Google Knowledge Graph >



- 객체: 5억 7천만개
- 사실: 180억개
- 지식추출 정확률@재현율: 45%@50%, 35%@80%

문제

- 계룡산의 **높이**는?
- 유성이 타버리지 않고 **땅에 떨어진 것은?**
- 자연어 QA 시스템을 평가하는 **방법**은?

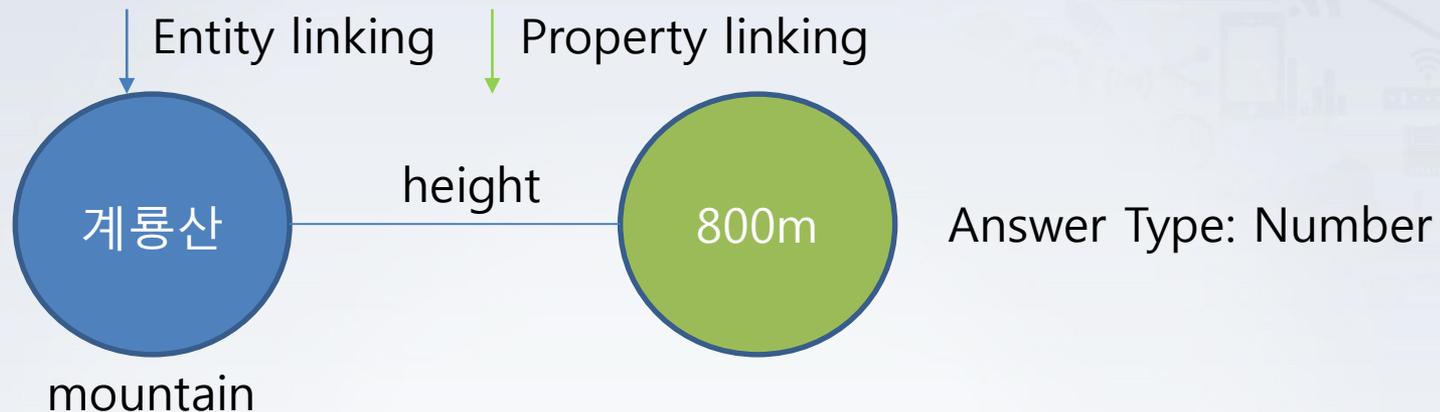
분석

- **모든 문제에 대한 명확한 분석 및 표현 불가**
- IBM 왓슨: 질문의 11%는 정답유형 결정 불가
- **모든 정답을 사전에 정형화된 지식으로 구축 불가**
- 지식: 선언적 지식, 절차적 지식, 경험적 지식 등

방법론

하이브리드 QA = Inductive QA + Deductive QA + Abductive QA

- Q: 계룡산[entity]의 높이[property]는?



- Conclusion must be true if all premises are true
- 도전 기술
 - 표현의 다양성, Partial parsing
 - 계룡산은 얼마나 높나요?
 - 계룡산의 높이는
 - Entity / Property disambiguation & linking
 - Ontology construction: 지식 표현 커버리지, 지식 구축 신뢰도

- Q: 오만원권 화폐에서 신사임당 옆에 그려져 있는 **과일**은?
 - Evidences
 - (+) 오만원권 앞에는 신사임당 옆에 포도가 그려져 있다.
 - (+) 오만원권에 있는 그림은 포도입니다.
 - (+) 포도는 오만원권 지폐에 신사임당 초상 뒤에 도안화되어 있다.
 - (-) 오만원권의 뒷면 그림으로는 매화와 대나무가 있다.
- Conclusion is probable based on supporting evidences
- 도전 기술
 - Linguistic knowledge(e.g., WordNet, FrameNet)
 - Uncertainty processing
 - 오만원권 화폐에서 신사임당 옆에 그려져 있는 **것**은?
 - Deep parsing: 어휘/문장/문맥의 이해
 - Paraphrasing

- inference to the best explanation
 - "a entails b" is used for inference
 - "snore" entails "sleep"
- Q: 김구는 언제 돌아가셨나요?
- 도전 기술: textual entailment in QA
 - Question entailment: 데이터 기반 표현 대중성
 - 김구는 언제 사망하셨나요?
 - Word/Phrase/Sentence entailment: "암살당하다" entails "사망하다"
 - 1949년 안두희에게 총격당하여 암살당하였다.



인간과의 의사소통을 뛰어넘어 지식소통이 가능하며,
전문가 수준 의사결정을 지원하는 “**인공두뇌**” 개발

도전
기술

(1) 자연어 이해 기술

자연어로 기술된 텍스트의 어휘, 문법, 문맥의 의미 이해
→ Wise Natural Language Understanding

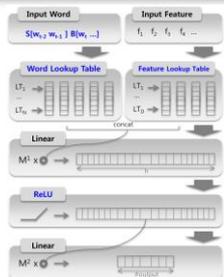
(2) 지식 학습 기술

빅데이터로부터 새로운 지식의 학습 및 생산
→ Wise KB

(3) 자연어 질의응답 기술

자연어 질문을 이해하고, 정형/비정형 지식으로부터 정답을 추론
→ Wise Question Answering

기술
독창성



Deep NLP Learning

$$L(e1 \stackrel{R}{\approx} e2)$$

$$L(e1 \stackrel{R}{\approx} e2) \Rightarrow K(e1 \stackrel{P}{\approx} e2)$$

Learning by Reading

Inductive QA
+
Deductive QA
+
Abductive QA
=

Hybrid Reasoning

발표 순서

I WiseQA 개요

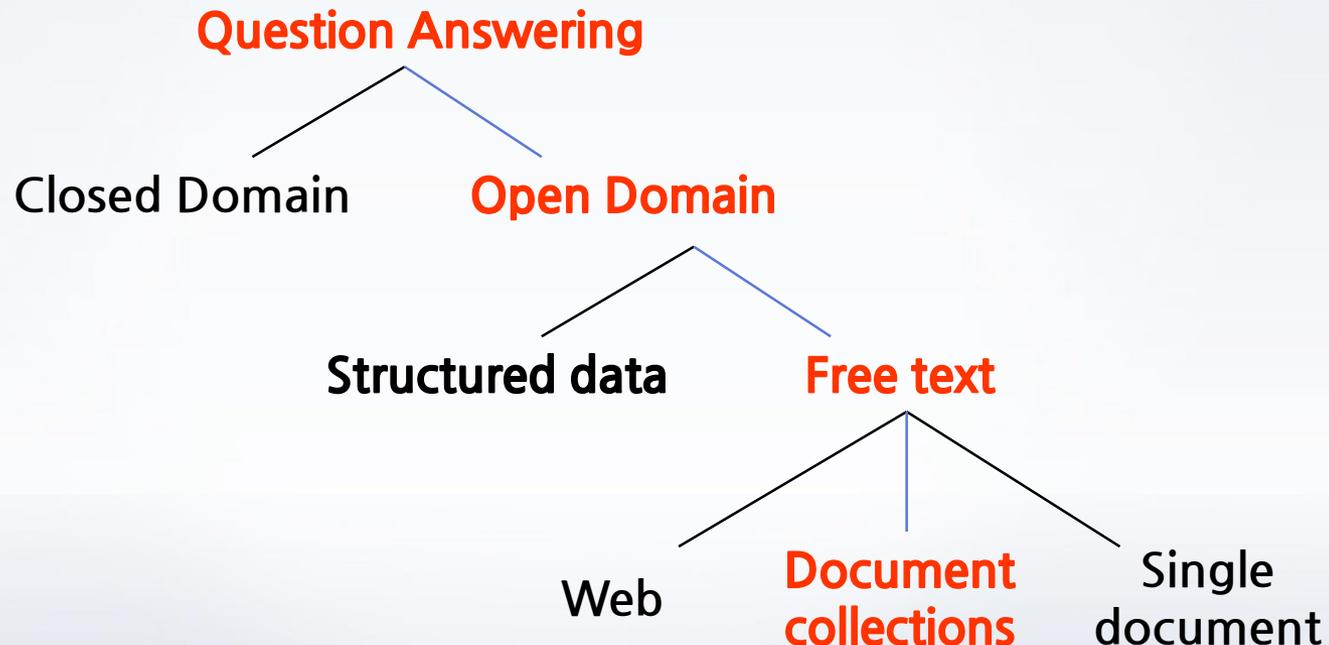
II QA 관련연구

III WiseQA 연구 성과 및 계획

IV 맺음말

Question Answering(QA)

- Find the answer to a question in a large collection of documents (By Bernardo Magnini)
 - **questions** (in place of **keyword-based query**)
 - **answers** (in place of **documents**)
- Research Context



Question Type & Answer Type

- **Question Type:** an idiomatic categorization of questions to distinguishing between different processing strategies and/or answer types
 - Factoid
 - List
 - Definition
 - Relationship
 - Superlative
 - Opinion
 - Command
 - etc
- **Answer Type:** the class of object or lexical type of sentence sought by the question
 - Exact answer: Person, Place, Date, Number, ...
 - Short answer: Definition, Explanation(Why, How)
 - Long answer: Summary, Report

- **Nature of information**
 - IR-based QA: useful in an open domain
 - Find Answers from a collection of documents
 - TREC QA, IBM Watson
 - Knowledge-based QA: useful in a close domain
 - Find Answers from a Knowledge Base
 - Google Knowledge Graph, Wolfram Alpha, Apple Siri
- **Nature of the technique**
 - Shallow linguistic methods
 - Simple, limited question & answer types
 - Deep linguistic methods
 - Complex, unlimited question & answer types

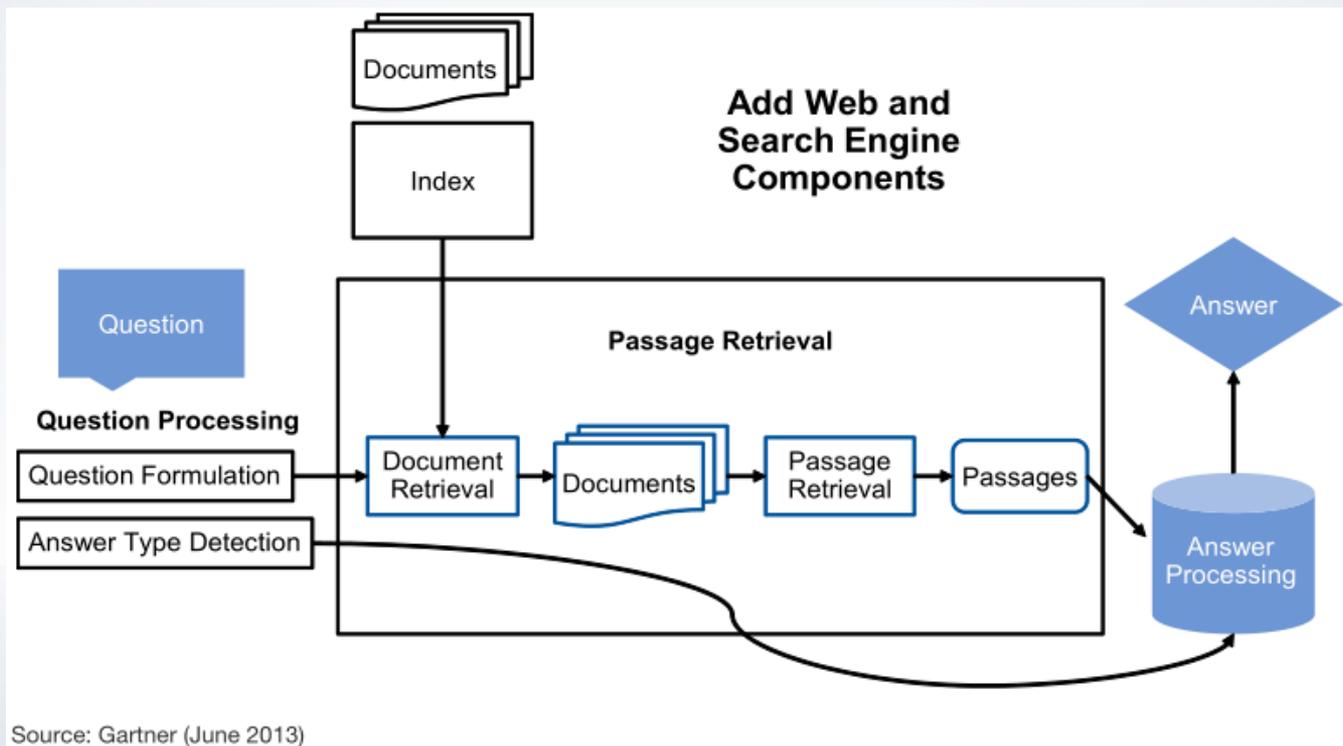
- TREC(Text Retrieval Conference) QA Track
 - 미국 NIST, ARDA, DARPA 후원
 - From 1999 ~ To 2007 (<http://trec.nist.gov/data/qamain.html>)
 - From 2002, the notion of **confidence was introduced**
 - 시스템이 제시한 1등이 정답인 경우만 고려함
 - Main task:
 - 자연어 질문(대부분 1문장)에 대한 Factoid, List, Definition 정답 찾기
 - 소스: 뉴스 텍스트(3GB)

• Type-and-Generate Method I

- Q: “**김구**는 언제 **태어**났나?” → Expected Answer Type: DATE
- A: “**김구**는 1876년 8월 29일(음력 7월 11일)에 황해도 해주 백운방 텃골(基洞)에서 (구)안동 김씨 김순영(金淳永, 당시 24세), 현풍곽씨 곽양식(郭陽植)의 딸 곽낙원 (당시 17세) 부부의 외동 아들로 **태어**났다.

(Wikipedia “김구’ 페이지에서 발췌)

- 정답후보: 1876년 8월 29일(음력 7월 11일), 1876년 8월 29일, 음력 7월 11일, 7월 11일



- Type-and-Generate Method I

1. Question Processing: 질문을 분석하여 찾고자 하는 정답유형 결정
2. Passage Retrieval: 질문에서 키워드 추출하여 관련된 구절 검색
3. Answer Processing: 검색된 구절을 분석하여 정답 추출 및 순위화

- 어려운 예 : 암살과 (총격당함→절명→죽다)의 의미 매칭

- Q : “김구는 언제 암살되었나?” → (Answer Type: Date)
- A : “[주어생략] 1949년 6월 26일, 12시 36분, 서울의 자택인 경교장에서 육군포병 소위 안두희에게 총격당하였다. [주어생략, 시간생략] 곧 병원으로 옮겨졌으나 절명하였다.” (Wikipedia “김구” 페이지에서 발췌)

• Type-and-Generate Method II (LCC)

1. 질문으로부터 찾고자 하는 Expected Answer Type 결정
2. 질문으로부터 키워드 추출 후, 분석하고자 하는 구절 검색
 - WordNet 활용 Lexical Alternation 대응 → Query reformulation
3. 질문 구조 and/or 정답유형 기반 검색결과 분석하여 정답추출
 - WordNet 활용 Semantic Alternation 분석

ANSWER TYPE	CATEGORY	NAMED ENTITY	CATEGORY
PERSON	→	person	
MONEY	→	money	
SPEED	→	price	
DURATION	→	quantity	
AMOUNT	→	number	

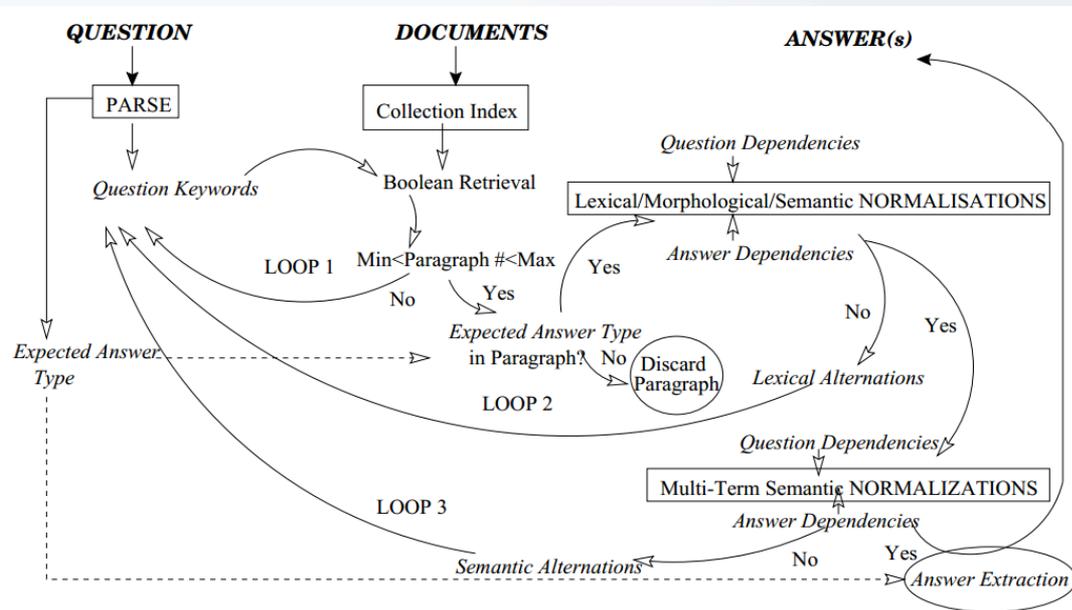


Figure 3: Retrieval Feedbacks in a Q/A System
(Source: S. Harabagiu, D. Moldovan)

219	Target: Iraqi defector Curveball
219.1	FACTOID What year did Curveball defect?
219.2	FACTOID What was Curveball's profession?
219.3	FACTOID What is Curveball's real name?
219.4	FACTOID Which intelligence service employed Curveball?
219.5	LIST Which US government officials accepted his claims regarding Iraqi weapons labs?
219.6	FACTOID Where does Curveball now live?
219.7	OTHER
254	Target: House of Chanel
254.1	FACTOID Who founded the House of Chanel?
254.2	FACTOID In what year was the company founded?
254.3	FACTOID Who is the president of the House of Chanel?
254.4	FACTOID Who took over the House of Chanel in 1983?
254.5	LIST What women have worn Chanel clothing to award ceremonies?
254.6	LIST What museums have displayed Chanel clothing?
254.7	FACTOID What Chanel creation is the top-selling fragrance in the world?
254.8	OTHER
269	Target: Pakistan earthquakes of October 2005
269.1	FACTOID On what date did this earthquake strike?
269.2	LIST What countries were affected by this earthquake?
269.3	FACTOID What was the final death toll from this earthquake?
269.4	FACTOID What was the strength of this earthquake?
269.5	FACTOID Where was the epicenter (latitude and longitude)?
269.6	LIST What countries supplied aid?
269.7	OTHER

Figure 1: Sample question series from the test set. Series 219 has a PERSON as the target, series 254 has an ORGANIZATION as the target, and series 269 has an EVENT as the target.

(Source: Overview of the TREC 2007 Question Answering Track)

TREC QA 성능

	2001	2003	2005	2006	2007
Tasks	Main/List/Context	Main/Passage	Main/Document/Relationship	Main/ciQA	Main/complex/ciQA
Document Collection	TIPSTER/TREC (979,000 documents; 3GB)	Corpus of English News (1,033,000 documents; 3GB)	Same as TREC 2003	Same as TREC 2003	AQUAINT-2 (News: 907,000 documents) + Blog06 corpus
# Questions	Main: 500 List:25 Context:42	Main: (Factoid: 413 List:37 Definition:50) Passage: 413	Main: (Factoid: 362 List: 93 Other:75) Document: 50 Relationship: 25	Main (Factoid: 403 List:89 Other:75) ciQA: 25	Main (Factoid: 360 List:85 Other:70) ciQA: 25
Question Source	MSNSearch and AskJeeves Logs	AOL and MSNSearch Logs	Same as TREC 2003	Same as TREC 2005	Same as TREC 2005
Correctness Judgments	Main/Context: Correct/ Incorrect/ Unsupported (Lenient: unsupported= correct; Strict: unsupported=incorrect) List: Correctness/Distinctness	Main (Factoid/list: Incorrect/Unsupported/Inexact/Correct) Definition: “Information nuggets” created and marked by assessors Passage: Incorrect/Unsupported/Correct	Main (Factoid/list: Incorrect/Unsupported/Inexact/Correct Other: same as TREC 2003 definition task) Document: relevant/not relevant Relationship: same as other task	Main (Factoid/list: Incorrect/Unsupported/Inexact/locally Correct/globally Correct Other: same as TREC 2005) ciQA: same as other task	Main (Factoid/list: Incorrect/Unsupported/Inexact/locally Correct/globally Correct Other: same as TREC 2005) ciQA: same as other task
Evaluation Measures	Main/Context: MRR List: Average Accuracy	Main: FinalScore= 0.5*FactoidScore+0.25*ListScore+0.25*DefScore Passage: Accuracy	Main: FinalScore =0.5*Factoid+0.25*List+0.25*Other Document: R-Prec, MAP Relationship:F($\beta=3$)	Main: FinalScore = 1/3*Factoid+1/3*List+1/3*Other ciQA: Pyramid F-Score	Main: FinalScore = 1/3*Factoid+1/3*List+1/3*Other ciQA: Pyramid F-Score
Best Main Task Results	0.68(MRR)	Final: 0.559 (Factoid:0.7, list:0.396, Def:0.442)	Final: 0.534 (Factoid:0.713, list:0.468, Other:0.248)	Final: 0.394 (Factoid:0.578, list:0.433, Other:0.250)	Final: 0.484 (Factoid:0.706, list:0.479, Other:0.281)

TREC QA → Type-and-generate QA

1. **Limited AT**: 질문으로부터 찾고자 하는 기정의 Answer Type 결정
2. 질문으로부터 키워드 추출 후, 분석하고자 하는 단락 검색
3. 질문 구조 및 정답유형 기반 검색결과 분석하여 정답추출

성능: 48.4% (Factoid: 70.6%, List: 47.9%, Other: 28.1%, 2007)

→ (-) **제한적 Answer Type** (Coarse grained: 6, Fine grained: 50, Li & Roth 2002)

Watson QA → Generate-and-type QA

1. **Unlimited AT**: 질문으로부터 찾고자 하는 Lexical Answer Type 동적 결정
2. 다양한 정답추출 방법 사용하여 정답후보 생성 및 정답후보의 타입 결정
3. 정답후보에 대한 근거분석 기반 최종 정답 합성 및 랭킹

성능: 70% (정확률), 86% (정확률) @ 70% (응답률) (2011)

→ (-) **제퍼디 퀴즈쇼의 정답의 95%는 위키피디아 표제어임**

POETS & POETRY: He was a bank clerk in the Yukon before he published “Songs of a Sourdough” in 1907.

Lexical Answer Type: 질문에서 정답으로 찾고자 하는 엔티티를 가르키는어휘

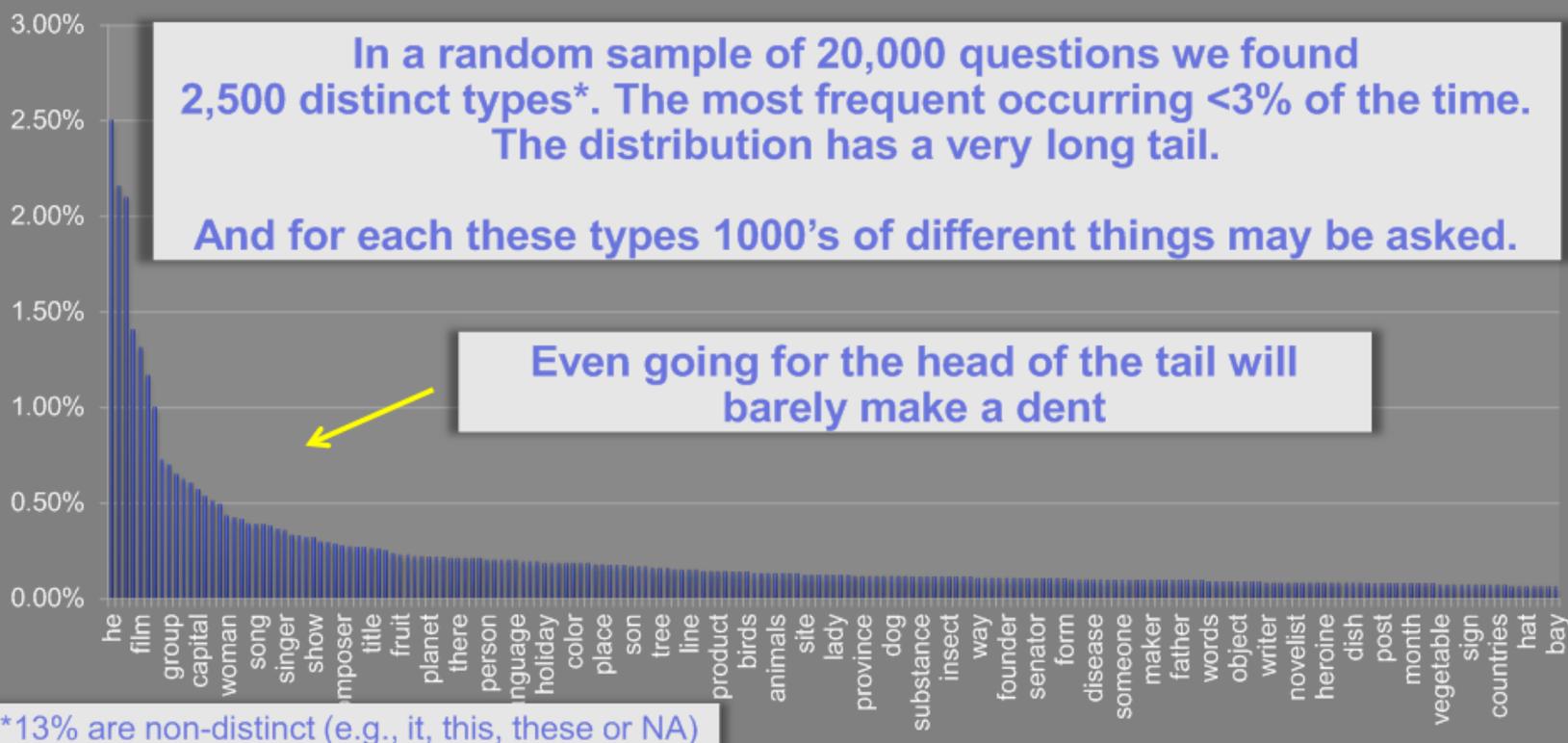
- “poet”, “he”, “clerk”

IBM Watson: Lexical Answer Types

Why LAT?

- **Open domain problem:**
 - Long tail answer type distribution
 - New answer types were being introduced at 12% to be roughly constant

We do NOT attempt to anticipate all questions and build specialized databases.

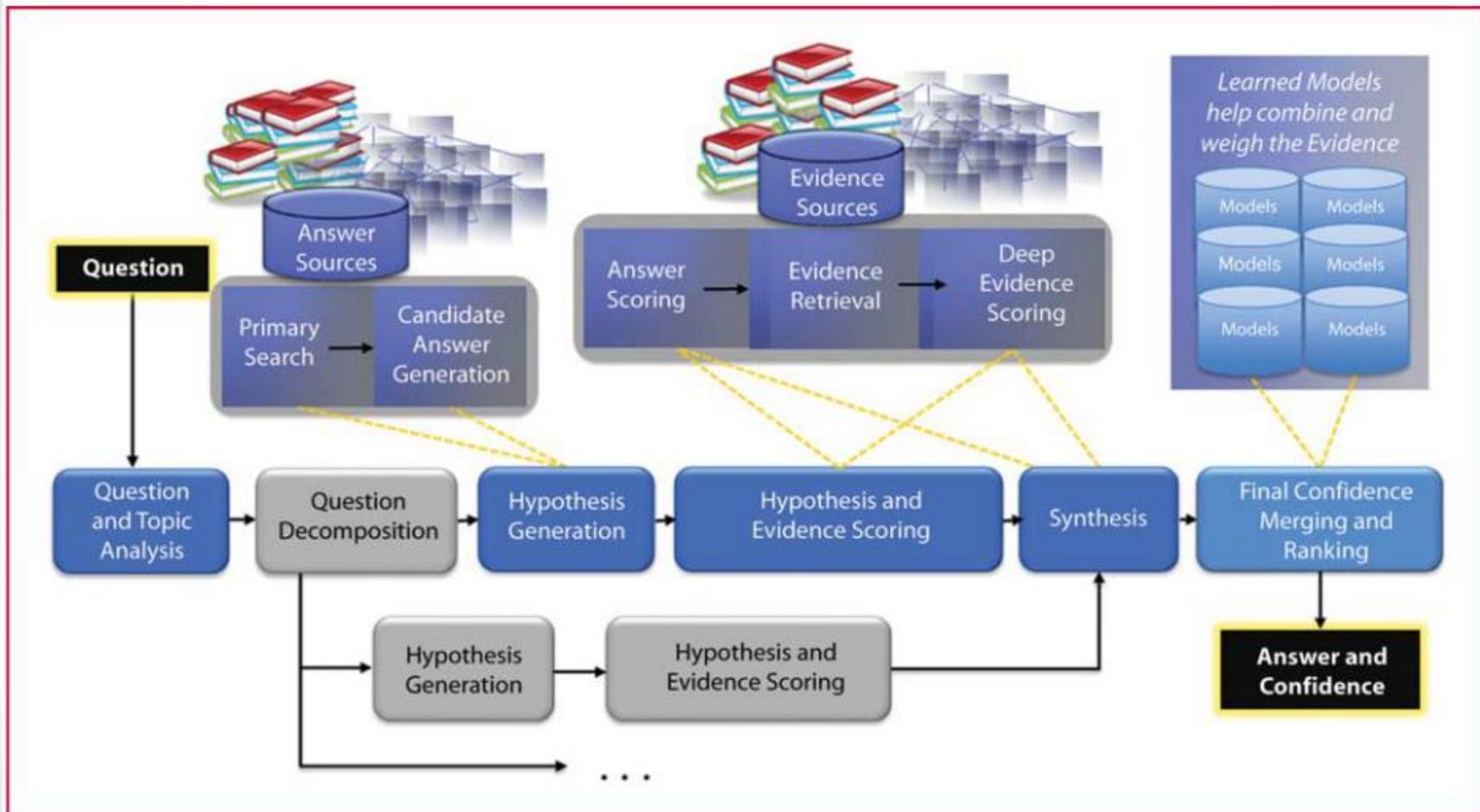


Our Focus is on reusable NLP technology for analyzing volumes of *as-is* text.
Structured sources (DBs and KBs) are used to help interpret the text.

IBM Watson: QA Architecture

핵심기술: Generate-and-type 방법 (Data-driven approach)

1. Hypothesis Generation: 질문을 분할하여 대량의 가설 생성
2. Evidence Retrieval & Scoring: 정답후보를 지지하는 근거 검색 및 점수 계산
3. Final Merging & Ranking: 질문 유형 별 최적화된 정답 랭킹



IBM Watson: Precision

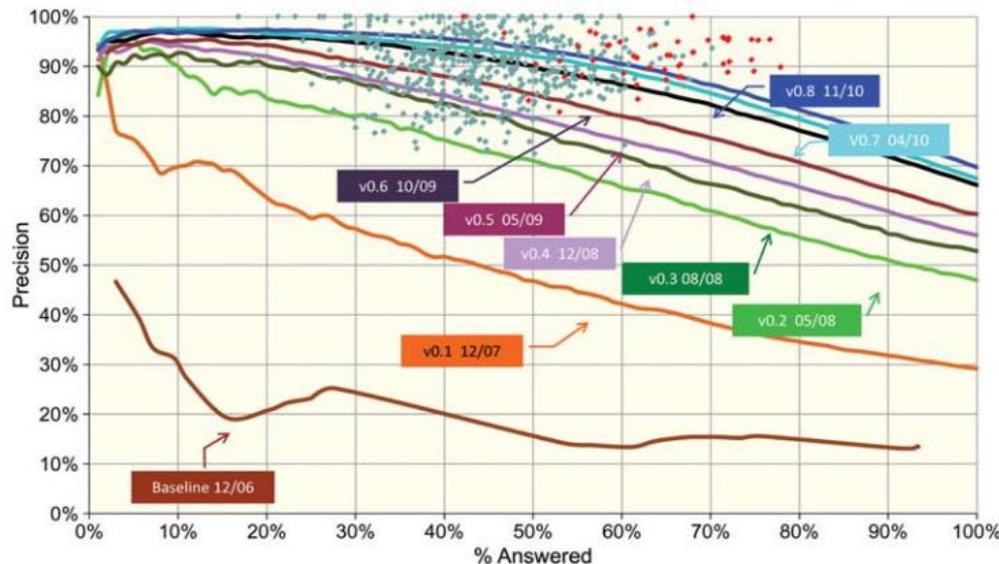
SW: Linux with UIMA, Hadoop, Lucene, Indri

HW: 90 x IBM Power 750 servers, 80 Teraflops (30억원)

- 2,880 CPU cores, 15 Terabytes of memory

IBM 왓슨이 우승하기까지...

- 2004년
 - IBM은 2nd 그랜드 챌리지로 '제퍼디' 퀴즈쇼 우승 목표 설정
- 2007년
 - 왓슨 Deep QA V0.1 개발
- 2011년
 - 왓슨 '제퍼디' 퀴즈쇼 우승(70%)
 - HW: 30억원 규모 슈퍼컴 사용
- 2012년
 - 왓슨 QA의 헬스케어, 빅데이터 분야 레퍼런스 SW 개발
- 2014년
 - 왓슨 응용 제품 개발을 위해 10억불 투자 결정



Deep QAs of Stanford, Facebook & MS

- WebQuestions dataset contains 5,810 question-answer pairs with common questions asked by web users (Berant et al. 2013)
- This dataset is built using Freebase as the KB by crawling questions

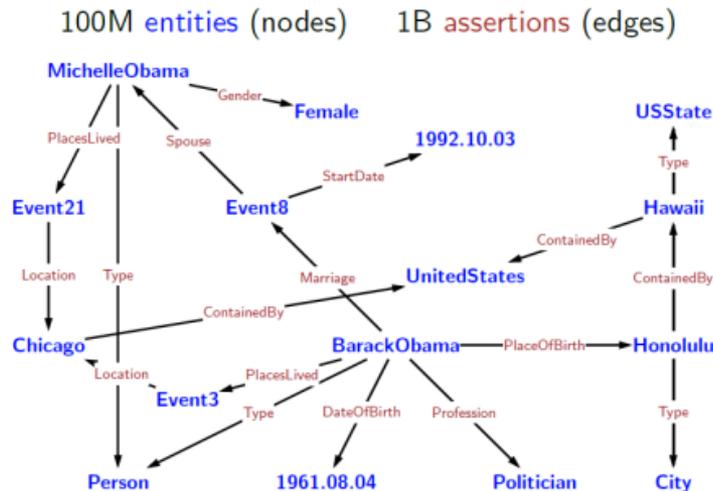


Figure 1. Freebase KB (Bollacker et al. 2008) has over 100m entities and the Stanford system used 41m of them for the WebQuestions benchmark.

Table 3. Results on the WebQuestions test set using Berant & Liang's F1 evaluation method. Other systems were reported by (Bordes et al 2014)

Method	F1 Score (Berant & Liang 2014)
Stanford (Berant et al. 2013)	31.4
Borders et al. 2013	29.7
Fader et al. 2014	35.0
Yao and Van Durme 2014	33.0
Stanford (Berant & Liang 2014)	39.9
Facebook (Bordes et al. 2014)	39.2
Ensemble of Facebook & Stanford (Bordes et al. 2014)	41.8
Our Approach: Microsoft Deep QA	45.3

발표 순서

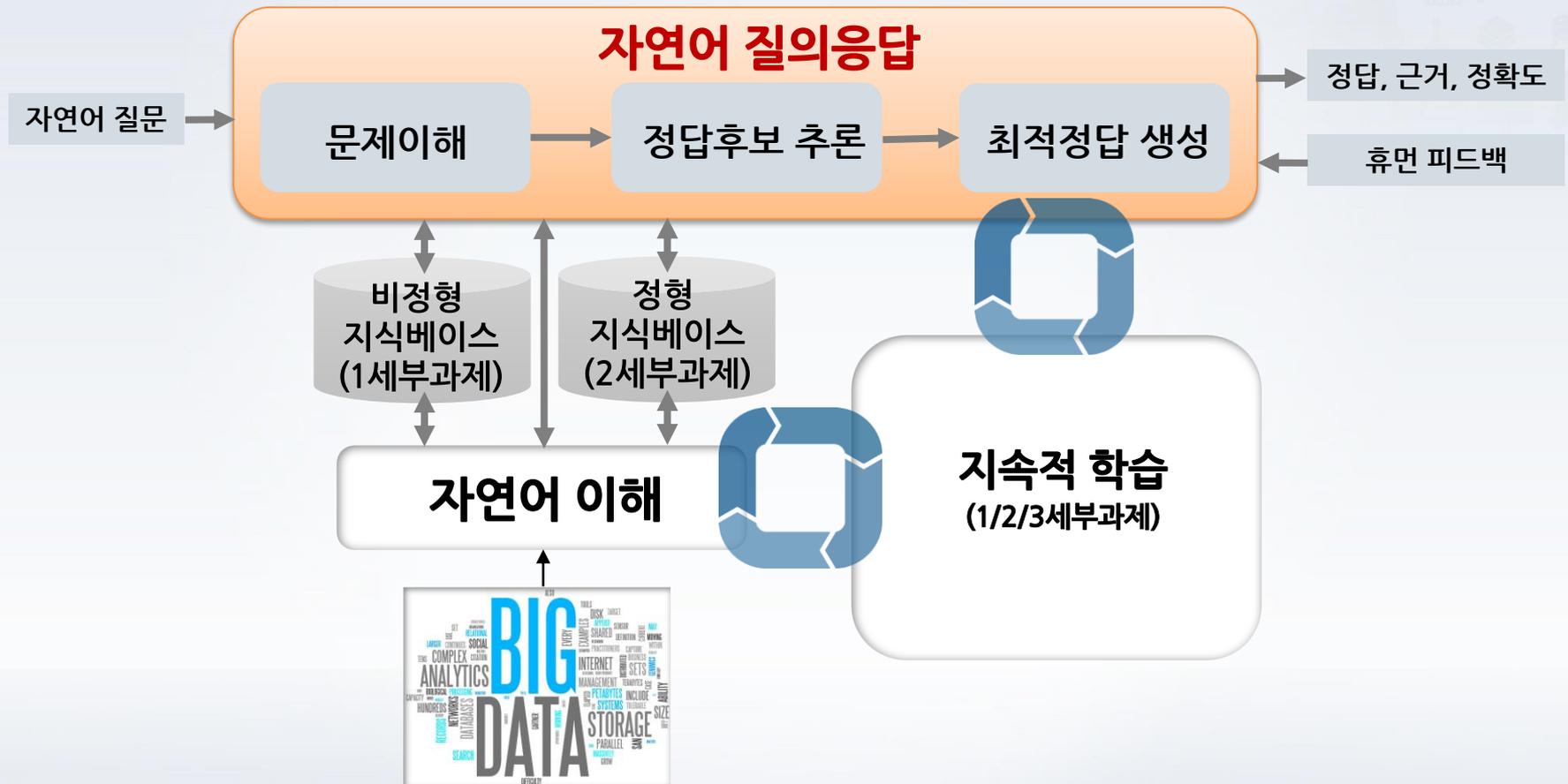
I WiseQA 개요

II QA 관련연구

III WiseQA 연구 성과 및 계획

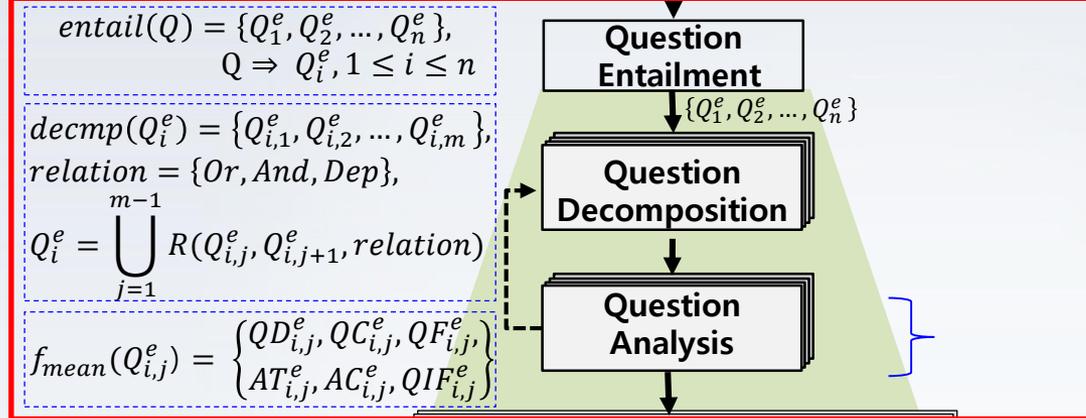
IV 맺음말

복잡한 자연어로 기술된 문제의 의미를 이해하고 정답을 추론하여 생성함



WiseQA Architecture

Natural Language Question Understanding



$$\text{entail}(Q) = \{Q_1^e, Q_2^e, \dots, Q_n^e\}, \\ Q \Rightarrow Q_i^e, 1 \leq i \leq n$$

$$\text{decmp}(Q_i^e) = \{Q_{i,1}^e, Q_{i,2}^e, \dots, Q_{i,m}^e\}, \\ \text{relation} = \{Or, And, Dep\}, \\ Q_i^e = \bigcup_{j=1}^{m-1} R(Q_{i,j}^e, Q_{i,j+1}^e, \text{relation})$$

$$f_{\text{mean}}(Q_{i,j}^e) = \begin{cases} QD_{i,j}^e, QC_{i,j}^e, QF_{i,j}^e \\ AT_{i,j}^e, AC_{i,j}^e, QIF_{i,j}^e \end{cases}$$

Answer Type
Axiom Proving

Answer Constraints
Axiom Proving

Answer, Confidence, Source

Q

남아메리카에 위치해 있으며, 수도가 카라카스인 이 나라의 국명은 '작은 베네치아'란 뜻을 가지고 있다



질문 분석 결과

<Question Entailment>

Q_1^e : 남미에 위치하고, 카라카스가 수도인 이 국가의 이름은 '작은 베네치아'라는 의미가 있다.

Q_1^e : 남아메리카에 위치하고, 국명이 '작은 베네치아'라는 의미가 있는 이 국가의 수도는 카라카스이다.

<Question Decomposition>

$Q_{1,1}^e$: 남미에 위치하고, 카라카스가 수도인 이 국가

$Q_{1,2}^e$: 이 국가의 이름은 '작은 베네치아'라는 의미가 있다.

$Q_{2,1}^e$: 남아메리카에 위치하고, 국명이 '작은 베네치아'라는 의미가 있는 이 국가

$Q_{2,2}^e$: 이 국가의 수도는 카라카스이다.

<Question Understanding Axioms>

$f_{\text{mean}}(Q_{1,1}^e) = \{$
 질문도메인(QD): 지리,
 질문클래스(QC): 단답형, 속성값요청형,
 질문 초점(QF): '이 국가',
 어휘정답유형(LAT): '국가',
 의미정답유형(SAT): LCP_COUNTRY,
 정답제약($LOC^{Explicit}$): 남미,
 정답제약($FRAME^{syntactic}$):
 (P: 수도, A0:X, A1:카리카스)

$\}, f_{\text{mean}}(Q_{1,2}^e), \dots, f_{\text{mean}}(Q_{2,2}^e)$

$R(f_{\text{mean}}(Q_{1,1}^e), f_{\text{mean}}(Q_{1,2}^e), And)$

$R(f_{\text{mean}}(Q_{2,1}^e), f_{\text{mean}}(Q_{2,2}^e), And)$

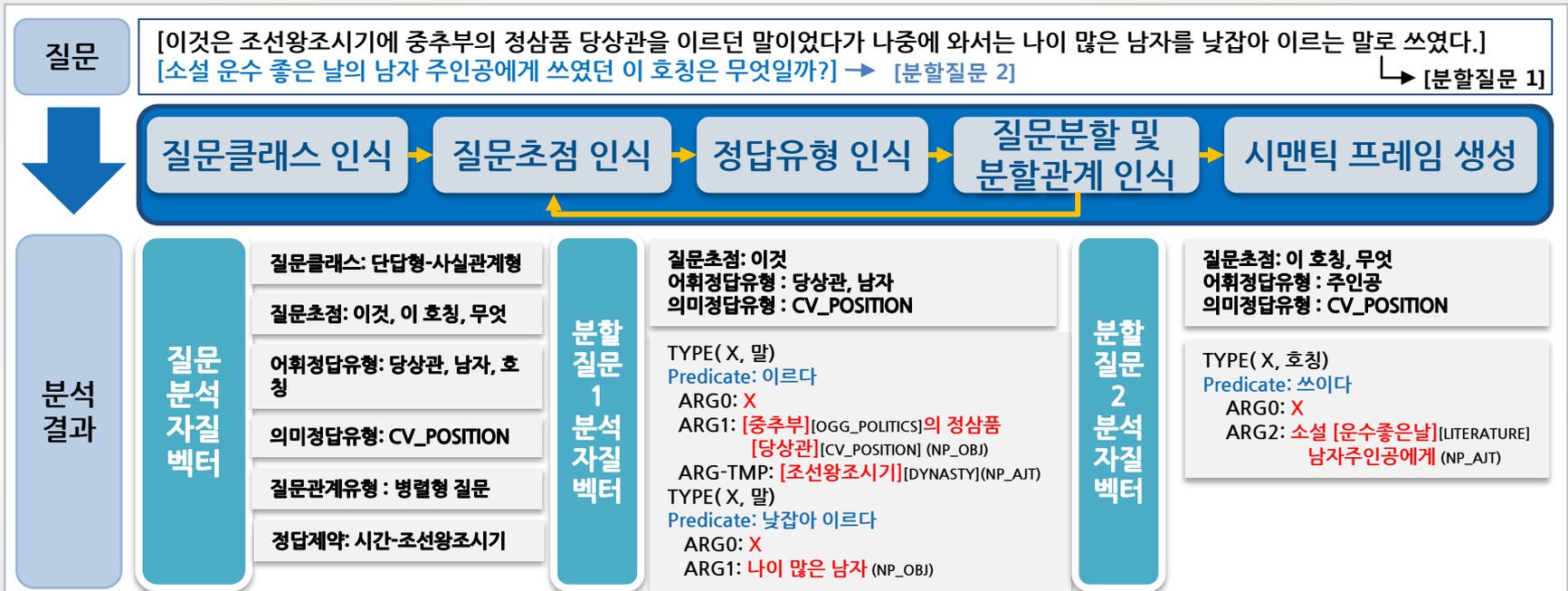
자연어 질문분석 기술

연구 목표

- 질문의 정답유형 인식 기술 개발
- 문제해결전략 생성을 위한 질문분할 기술 개발
- 지식베이스 검색을 위한 질문 시맨틱 프레임 생성

주요 성과

- **Hybrid 정답유형 인식** 및 WiseWordNet 기반 정답유형 상호검증 기술 개발
 - 기계학습과 규칙방법을 결합한 어휘정답유형(LAT)과 의미정답유형(SAT)의 Hybrid 정답유형 인식 및 검증
- 질문초점을 이용한 규칙기반 **질문분할 및 분할관계 인식** 기술 개발
- 심층언어 분석 및 패턴 기반 **질문 시맨틱 프레임 생성** 기술 개발



자연어 질문 분류

분류기준	대분류	중분류	소분류	설명	
정답 형태	단답형			단답형식의 명사(구) 또는 어휘로 정답을 제시해야 하는 경우	
	서술형			주관식 문장이나 개조식으로 정답을 제시하는 경우	
	나열형			정답이 1개 이상인 형태의 질문	
	순서형			정답이 순서대로 제시되어야 하는 질문	
	짝짓기형			속성과 정답이 짝을 이루어 제시되어야 하는 경우	
질문 형태	정의형	용어요청형		정의문이 제시되고 용어를 찾는 문제	
		의미요청형		용어가 제시되고 정의문을 찾아 정답을 제시하는 문제	
	사실관계형	속성값요청형		속성을 제시하고 속성값을 찾는 문제	
		속성요청형		속성값이 제시되고 속성을 찾는 문제	
	추론형	논리추론형	연상형		다양한 질문/용어 등이 공통적으로 연관되는 정답을 추론하는 질문
			관계추론형		Entity나 Event의 속성이 제시되지 않고, 단지 관련된 속성값을 찾는 질문
		연산추론형	속성비교형		속성값에 대한 비교 연산이 수행되어야 하는 질문
			시간비교형		속성값 중, 시간정보에 대한 연산이 수행되어야 하는 질문
			계산형		속성값에 대한 사칙연산이 수행되어야 정답을 제시할 수 있는 질문
	OOD형			멀티미디어 정보와 함께 제시되는 질문	
질문 구조 형태	일반형			괄호채우기를 제외한 주관식 문제	
	괄호채우기형			심볼 등을 통해 Blank를 채워넣는 형태의 질문	
	다지선다형			제시된 보기들 중, 하나를 선택하는 질문 (긍정질문만)	
	부정형	다중선택형			부정형을 긍정형으로 변환 시, 보기들 중 올바른 것을 복수개 선택하는 질문의 형태로 됨
		뮴음형			부정형을 긍정형으로 변환 시, 보기들 중 유사한 특징으로 묶는 형태의 질문

WiseQA 평가 질문(장학퀴즈 문제)

이수익은 이별의 비애를 담은 그의 시에 <우울한 '이것'>이라는 제목을 붙여 쓸쓸하고 감상적인 느낌을 더했다. 프랑스 대중이 즐기는 가요를 지칭하는 이것은 무엇일까?	샹송
이 별자리는 황도 12궁의 마지막 별자리로 미의 여신 아프로디테와 그녀의 아들 에로스가 변한 것이라고 전해진다. 2월에서 3월의 탄생 별자리이기도 한 이 별자리는 무엇일까?	물고기자리
이 단어는 다른 나라 항구에 상업용 선박이 드나드는 것을 법으로 금지하는 것을 의미하는 무역용어이다. 언론에서는 특정 시점까지 보도를 유예하는 것을 뜻하는 말로 쓰이는데 이 단어는 무엇일까?	엠바고
이것은 조선 시기 흥선대원군이 경복궁 중건을 위해 강제로 거둔 기부금이다. '백성들이 원망하며 내는 세금'이라는 별명을 얻기까지 한 이것은 무엇일까?	원납전
프랑스 화가인 이 사람은 원색을 이용해 강렬하면서도 개성적인 작품을 그렸다. 야수파 운동을 주도했으며 <붉은 방>을 그린 이 사람은 누구일까?	마티스
국민에게 인기가 있고 명망 있는 인물을 후보로 영입하는 데 유리한 제도로서, 정당에 관계없이 누구나 참여해 당의 대선 후보를 선출하는 이 제도는 무엇일까?	오픈 프라이머리
이것은 조선 시기 정조가 즉위한 후 기존의 호위 기구인 숙위소를 대신해 설치한 국왕의 친위 부대이다. 수도에 내영을 두고 수원에 외영을 설치했던 이 부대는 무엇일까?	장용영
이것은 날지 못하는 새로, 조류 가운데 드물게 사람을 겁내어 피하지 않는다. 남반구의 해양 환경에 주로 서식하며 땅 위를 뒤뚱거리며 걸어 다니는데 이 새는 무엇일까?	펭귄
이 열매는 지중해 연안에서 재배되는 것이 가장 품질이 좋으며, 프랑스 남부에 있는 망통에서는 매년 2월에 이 열매를 이용한 축제가 열린다. 신맛을 내는 이 열매는 무엇일까?	레몬
이들은 중세 시대 교회에서 여성 성악가를 쓸 수 없었기 때문에 생겨났다. 여성과 비슷한 성질의 목소리를 가진 남성 성악가를 말하는 이 용어는 무엇일까?	카스트라토

비정형 정답후보 추론 기술

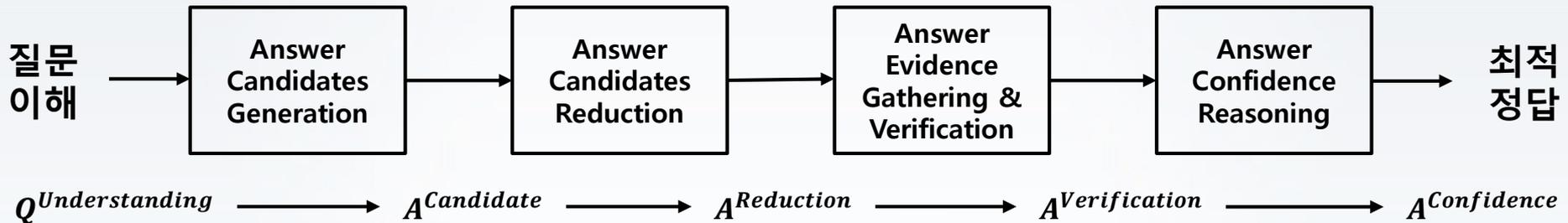
연구목표

- 텍스트 빅데이터 대상 정답후보 색인/검색
- 오픈 도메인 질의응답을 위한 정답후보 생성
- 심층분석 기반 정답가설 근거 검색 및 검증
- 정답의 다양한 근거 기반 정답가설 통합 및 순위화

주요성과

- **다중모델** 기반 정답후보 색인 및 검색
- 정답가설 생성을 위한 **정답후보 유형 추론 및 제약** 기술 개발
- **문맥 정보 유사도** 기반 정답가설 근거 추론
- 정답가설의 **다중 근거 통합 및 LTR 기반 순위화**

* LTR(Learning to Rank)



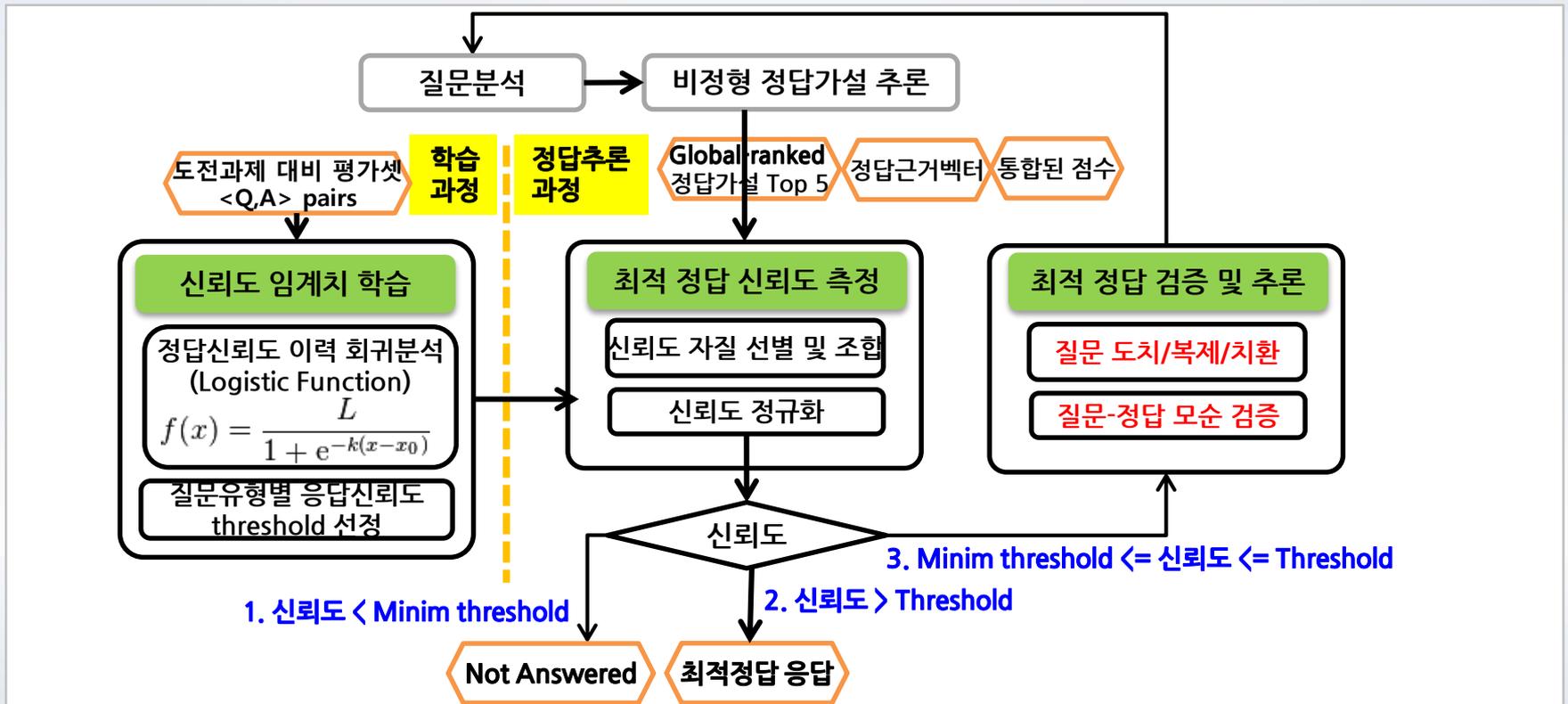
정답 신뢰도 추론 기술

연구목표

- 다중근거 기반 정답가설의 신뢰도 측정
- 응답전략 기반 최적 정답 신뢰도 임계치 학습
- 정답가설 검증을 통한 최적 정답 추론 방법 설계

주요성과

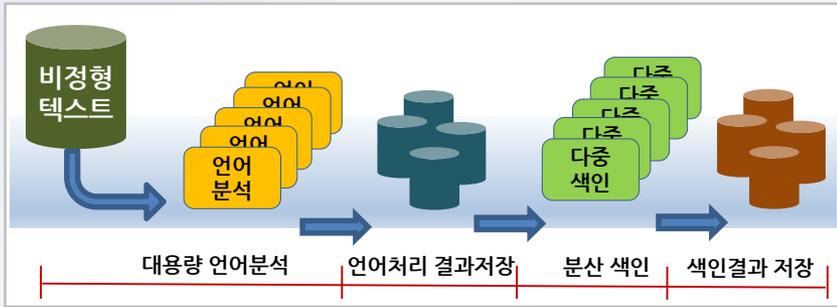
- 정답 유형별 신뢰도 자질 선별 및 조합의 학습
- 시뮬레이션을 통한 최적 정답 신뢰도의 임계치 학습 방법 고안
- 질문도치, 질문복제, 질문치환을 통한 최적 정답 검증 방법 도출



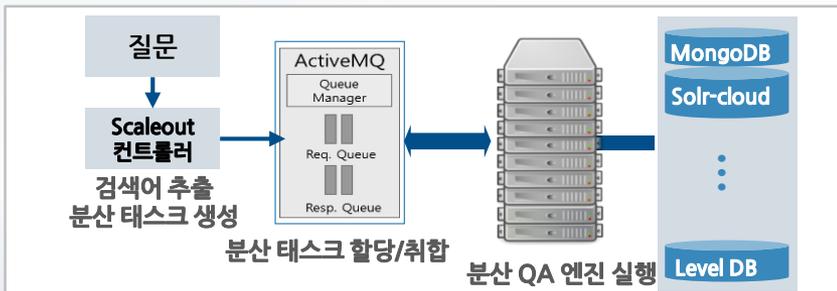
WiseQA 분산병렬처리 플랫폼 개발

연구목표

- 대용량 텍스트의 언어 분석/색인/저장 기술 개발
- 빠르고 효율적인 QA 연산을 위한 실시간 처리 WiseQA 플랫폼 개발
- * IBM 왓슨은 30억원 규모의 슈퍼컴퓨터 사용



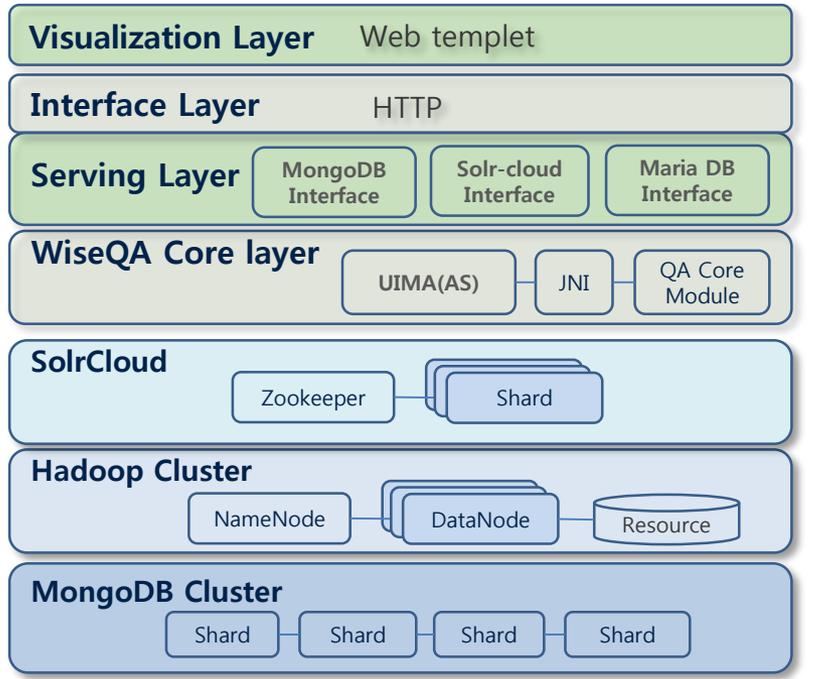
〈대용량 언어분석/색인/저장 기술 연구 목표〉



〈WiseQA 분산병렬처리 플랫폼 개발 기술 연구 목표〉

주요성과

- 맵리듀스 기반 **대용량 언어 분석/색인/저장 기술** 개발
- PC 클러스터링을 통한 **빠르고 저렴한 WiseQA 플랫폼 개발**
 - * 효율적 다중모델 검색을 위한 **분산검색 기술** 개발
 - * UIMA-AS 기반 **실시간 WiseQA 플랫폼 개발**



〈WiseQA 분산병렬처리 플랫폼 개발 기술 구성도〉

- Deep NLP parsing: Rule + Supervised learning + DL
 - SRL, CoRef. Resolution, Zero-anaphora resolution
 - Textual entailment: DL using WiseWordNet
- NL Question Understanding for axiom proving
- Tight coupling of IR-based & KB-based QA
 - Natural language to logic representation(1세부)
 - logic representation to structured query language(2세부)
- Hybrid reasoning QA
 - Deductive + Inductive + Abductive approaches to learn best combination of approaches
- Answer confidence reasoning

발표 순서

I WiseQA 개요

II WiseQA 연구성과

III WiseQA 연구계획

IV 맺음말

Factoids

simple facts that can be extracted from short text strings

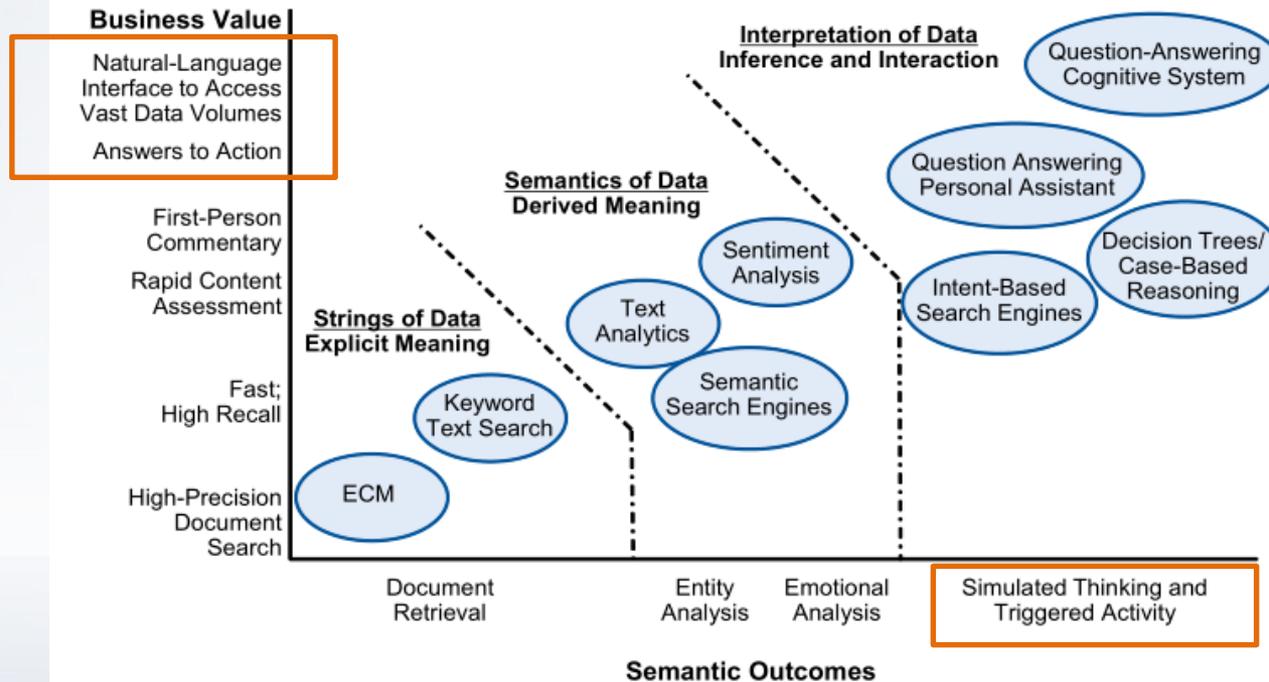
Simple Summarization (Smart Advisor)

a shortened version of one or more documents or content objects

Focused Summarization (Decision Supporter)

interpret multiple information sources and then stitch together a coherent answer

Figure 2. The Information Retrieval Family



Source: Gartner (June 2013)

- 2001: A Space Odyssey (1968년)

- 아서 C. 클라크's three laws
 - 어떤 노년의 과학자가 무엇이 가능하리라고 한다면 그것은 거의 확실히 맞다. 그러나 그가 무엇이 불가능하리라고 한다면 틀릴 가능성이 높다.
 - 가능성의 한계를 발견하는 유일한 방법은 불가능할 때까지 시도해 보는 방법밖에 없다.
 - 충분히 발달한 과학은 마법과 구분할 수 없다.



감사합니다.