

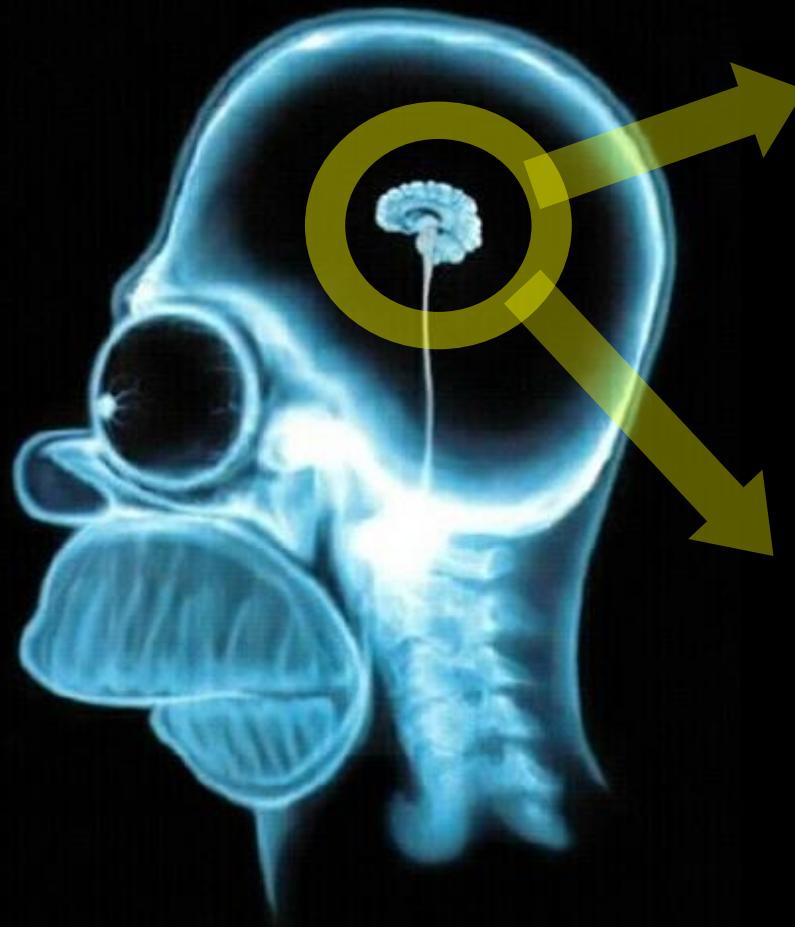
AI Applications 워크샵 – 생활속의 인공지능
인공지능연구회 / 한국정보과학회 / 네이버 그린팩토리 Connect Hall / 2016.6.3

초연결 지식과 인공지능

이 경 일 / Saltlux, Inc.

Act One

Hyper-Connection



Neuron

~100 Billion #
~ # of
Web Pages

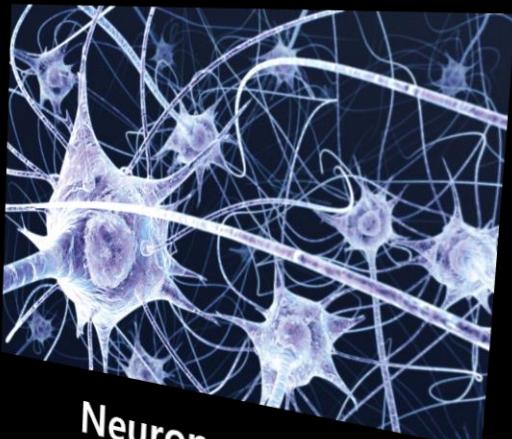


Synapse

~100 Trillion #
~ # of
Web Links



X 1,300

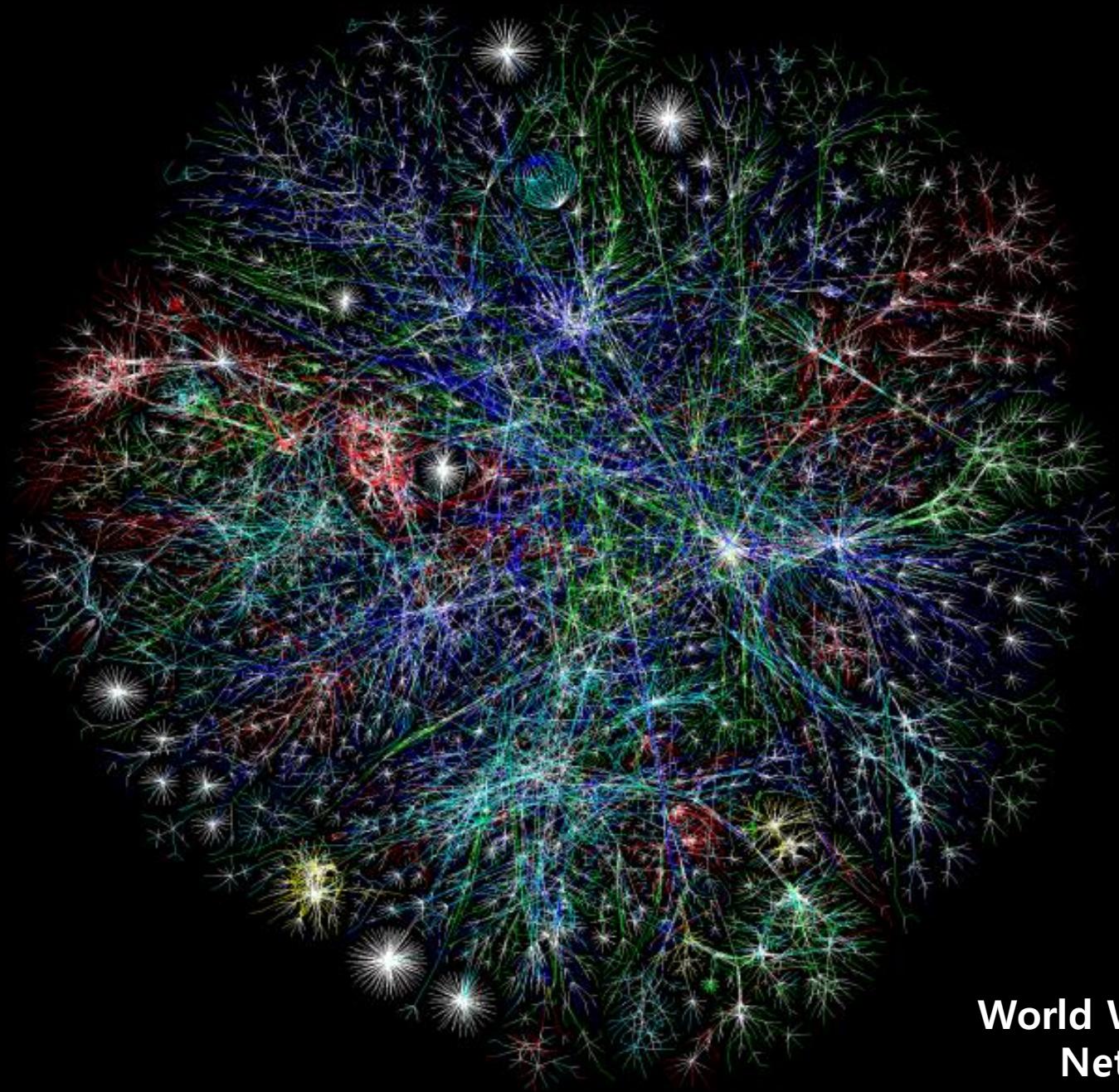


Neurons : 100B #

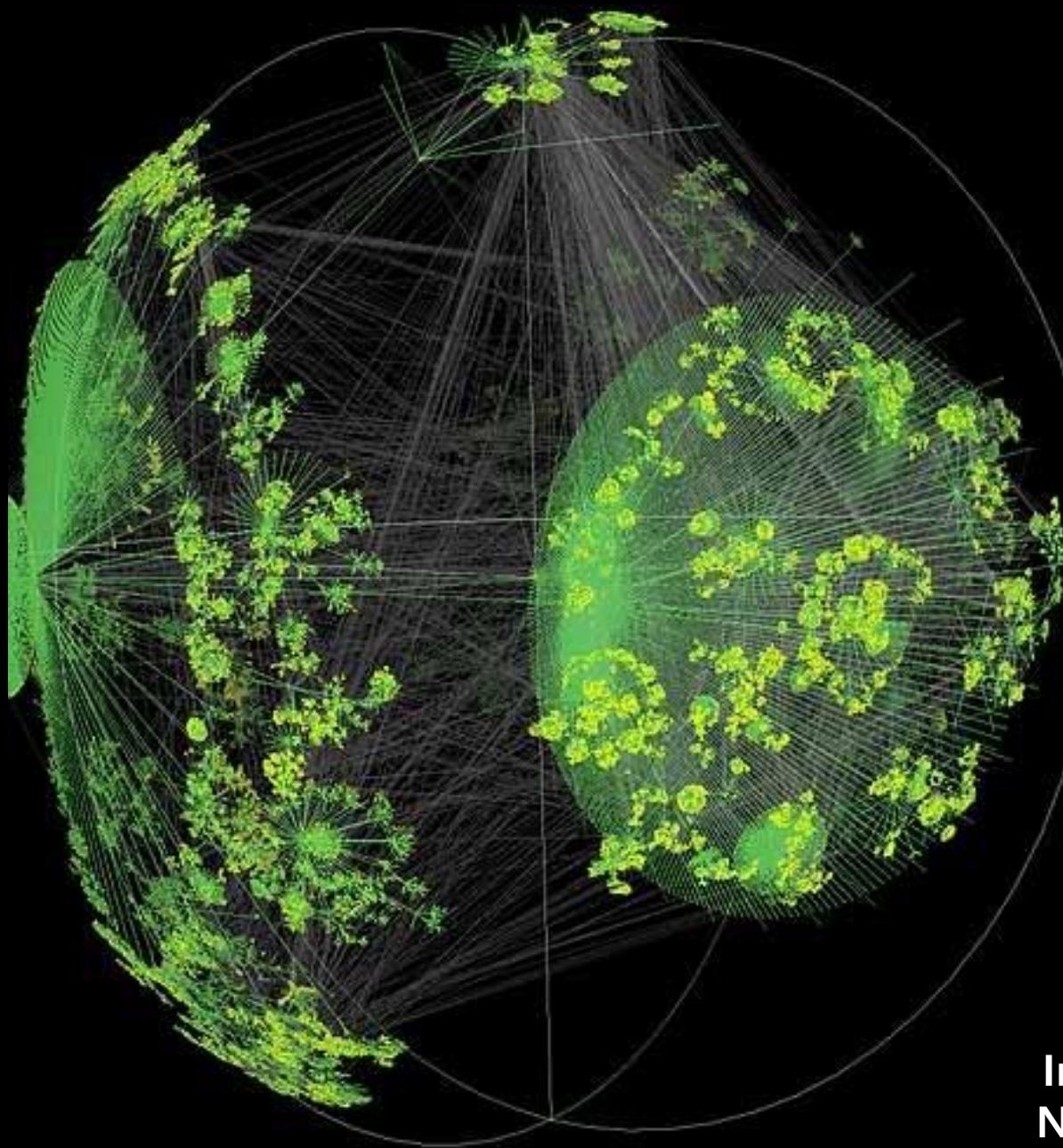
1 : 1000



Synapse : ~100T #



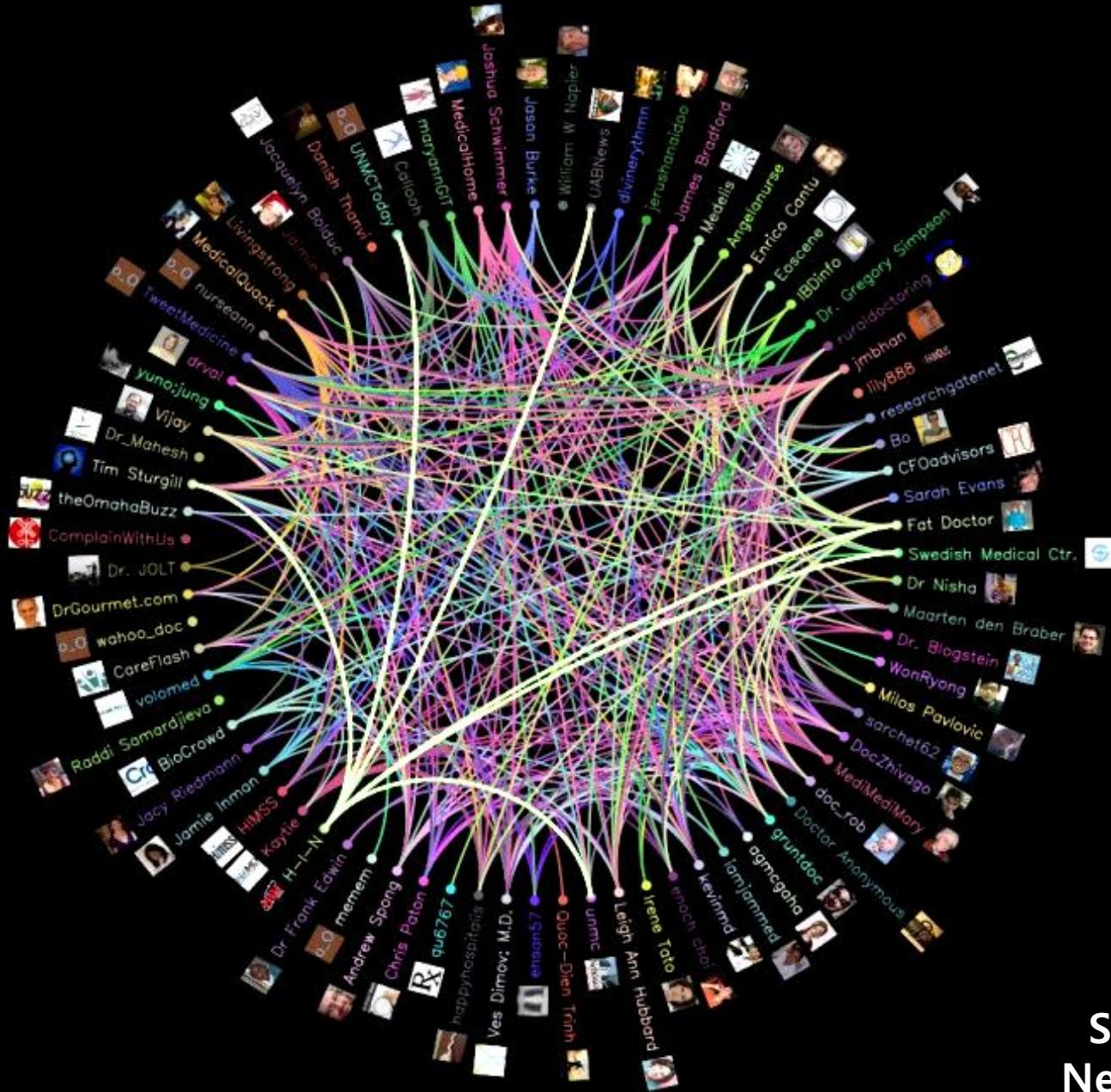
**World Wide Web
Network**



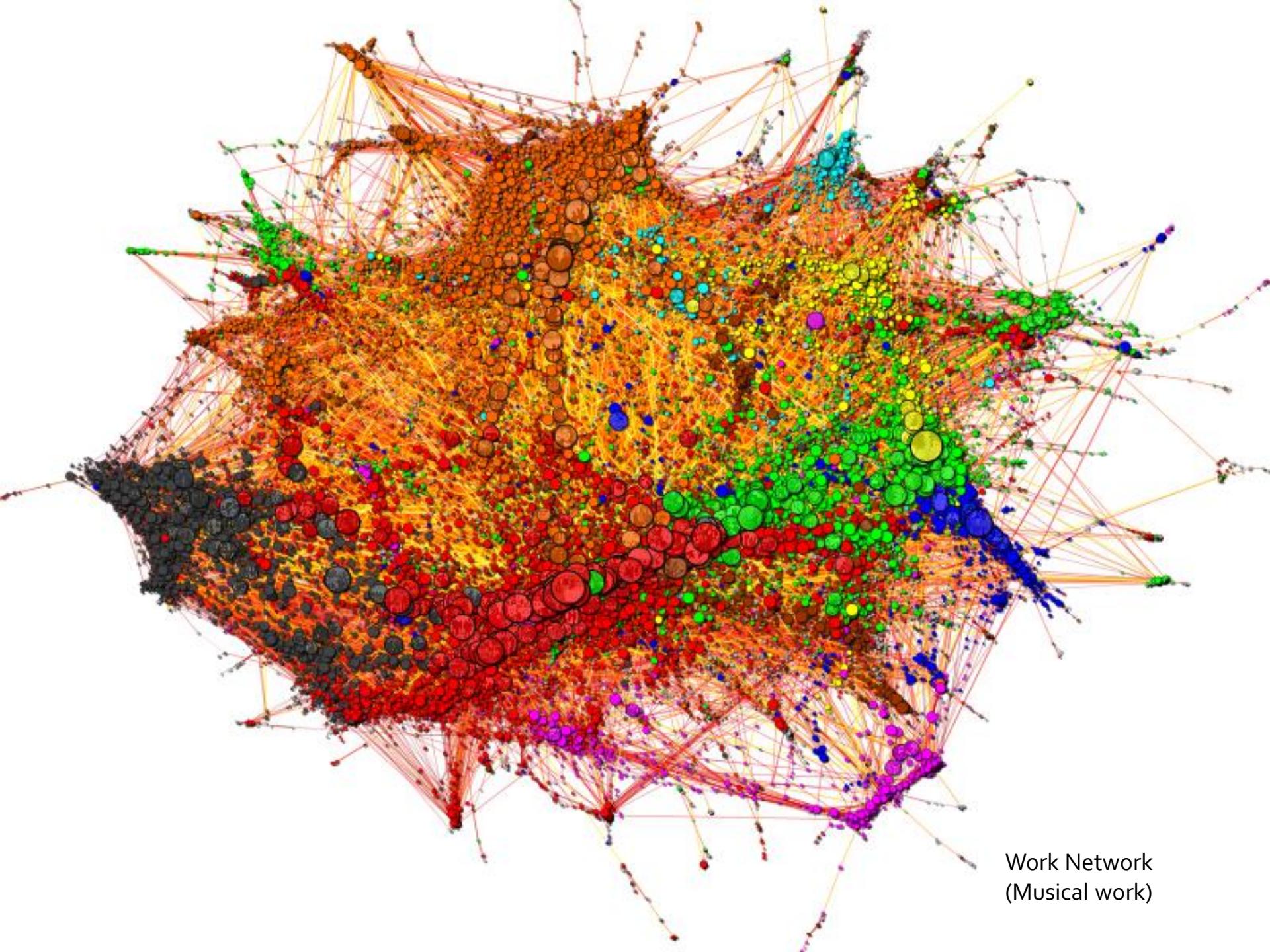
**Internet
Network**



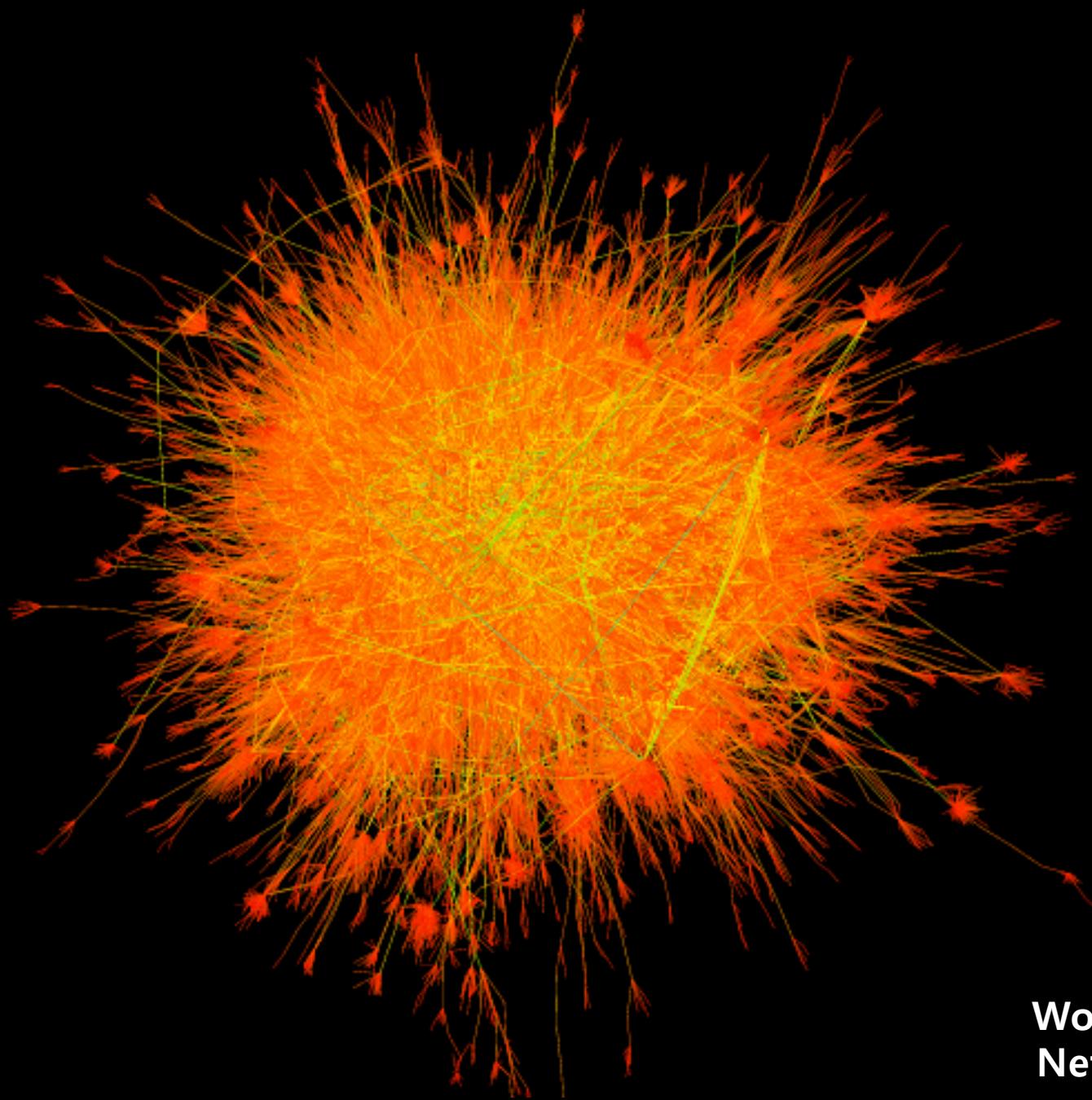
**Mobile
Network**



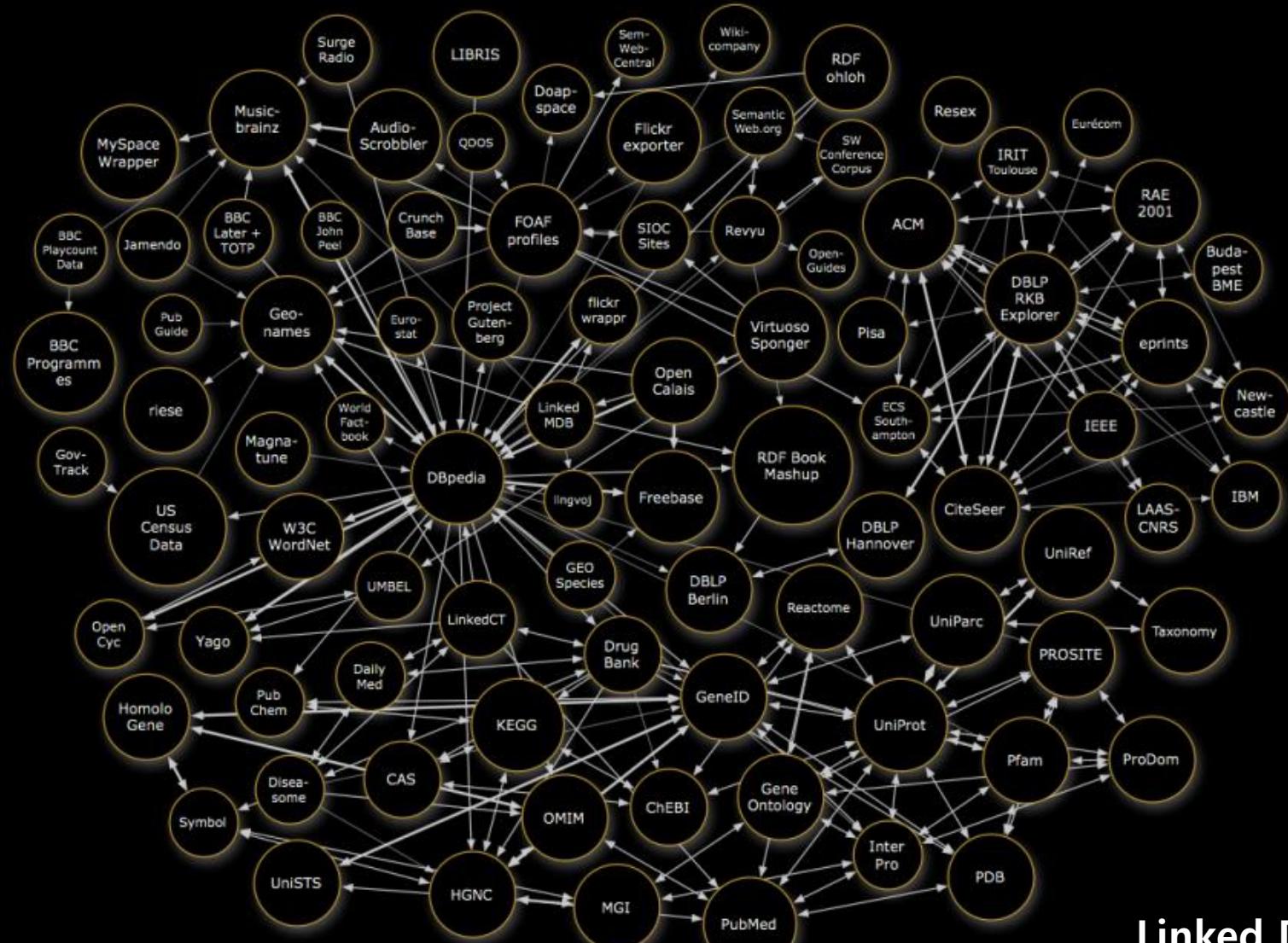
Social Network



Work Network
(Musical work)

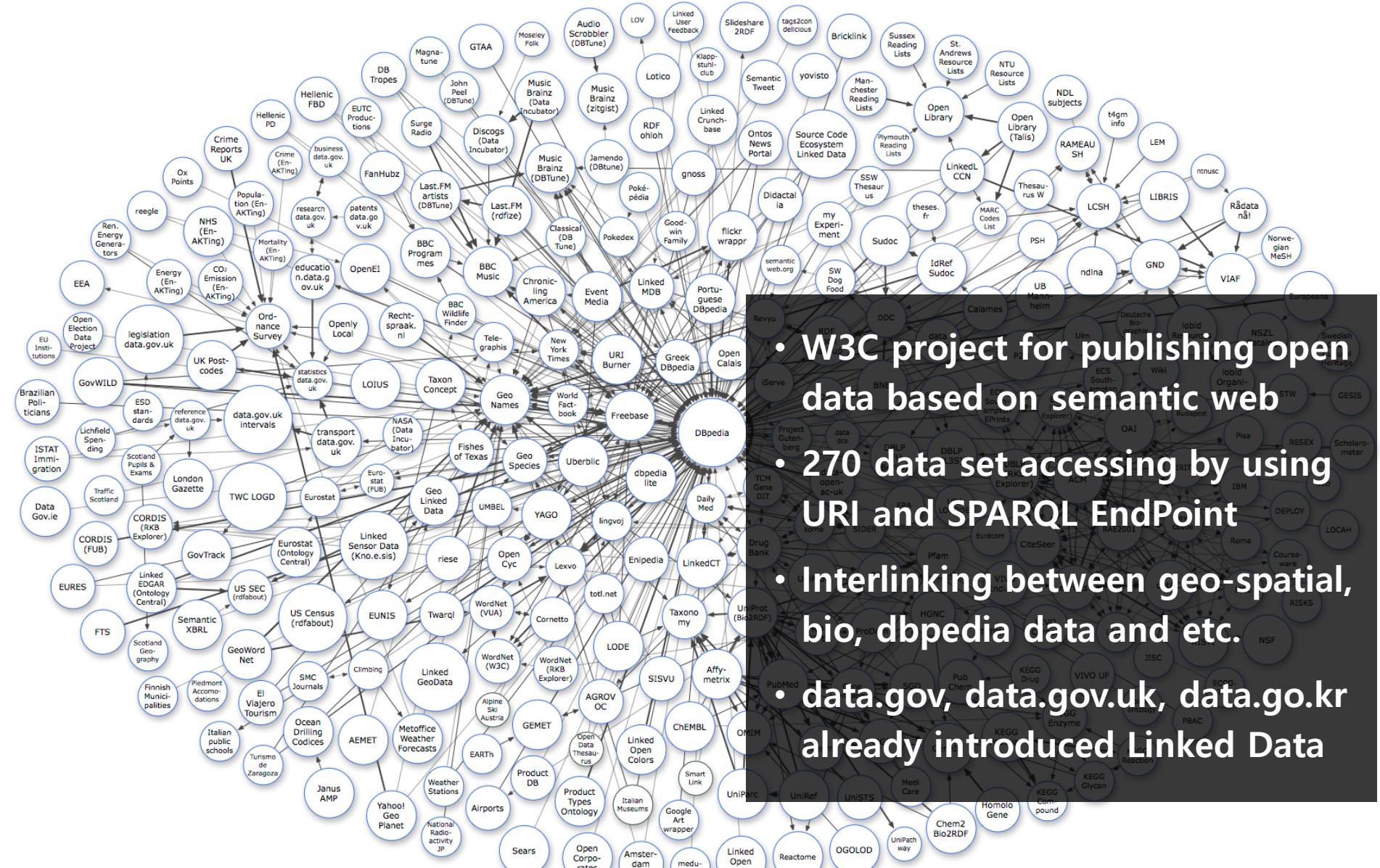


**WordNet
Network**



Linked Data Network

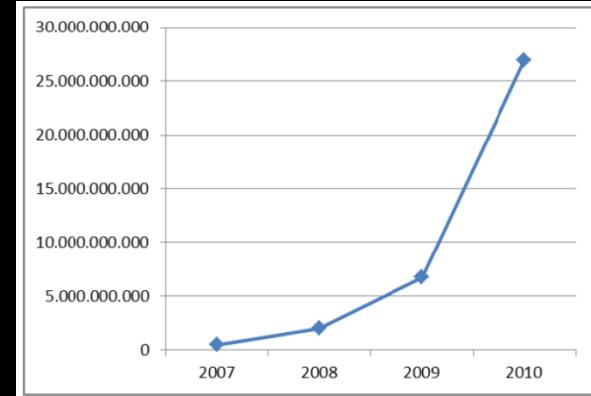
LOD project (Linking Open Data)



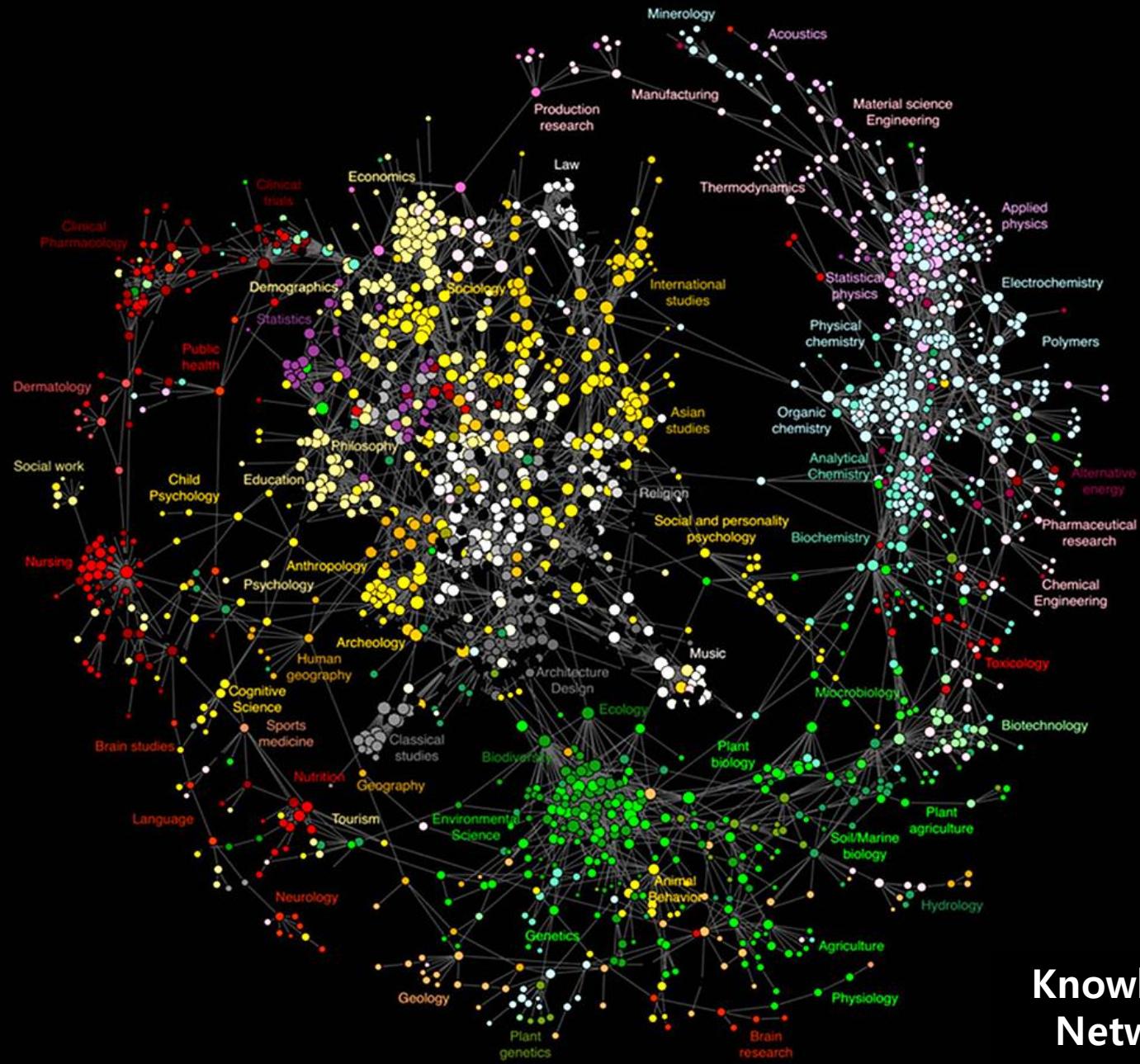
- W3C project for publishing open data based on semantic web
- 270 data set accessing by using URI and SPARQL EndPoint
- Interlinking between geo-spatial, bio, dbpedia data and etc.
- data.gov, data.gov.uk, data.go.kr already introduced Linked Data

LOD project Statistics

Year	Datasets	Triples	Growth
2007	12	500.000.000	
2008	45	2.000.000.000	300%
2009	95	6.726.000.000	236%
2010	203	26.930.509.703	300%



Domain	Data Sets	Triples	Percent	RDF Links	Percent
Cross-domain	20	1,999,085,950	7.42	29,105,638	7.36
Geographic	16	5,904,980,833	21.93	16,589,086	4.19
Government	25	11,613,525,437	43.12	17,658,869	4.46
Media	26	2,453,898,811	9.11	50,374,304	12.74
Libraries	67	2,237,435,732	8.31	77,951,898	19.71
Life sciences	42	2,664,119,184	9.89	200,417,873	50.67
User Content	7	57,463,756	0.21	3,402,228	0.86
	203	26,930,509,703		395,499,896	



Knowledge Network

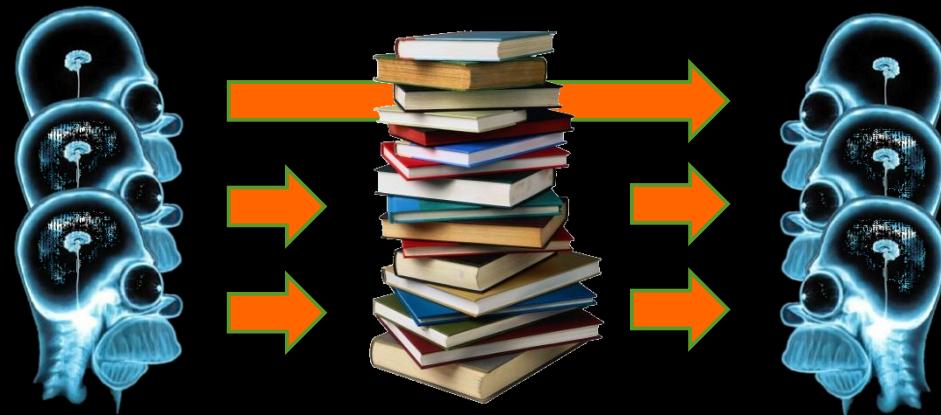
Intellectual Activity of Human



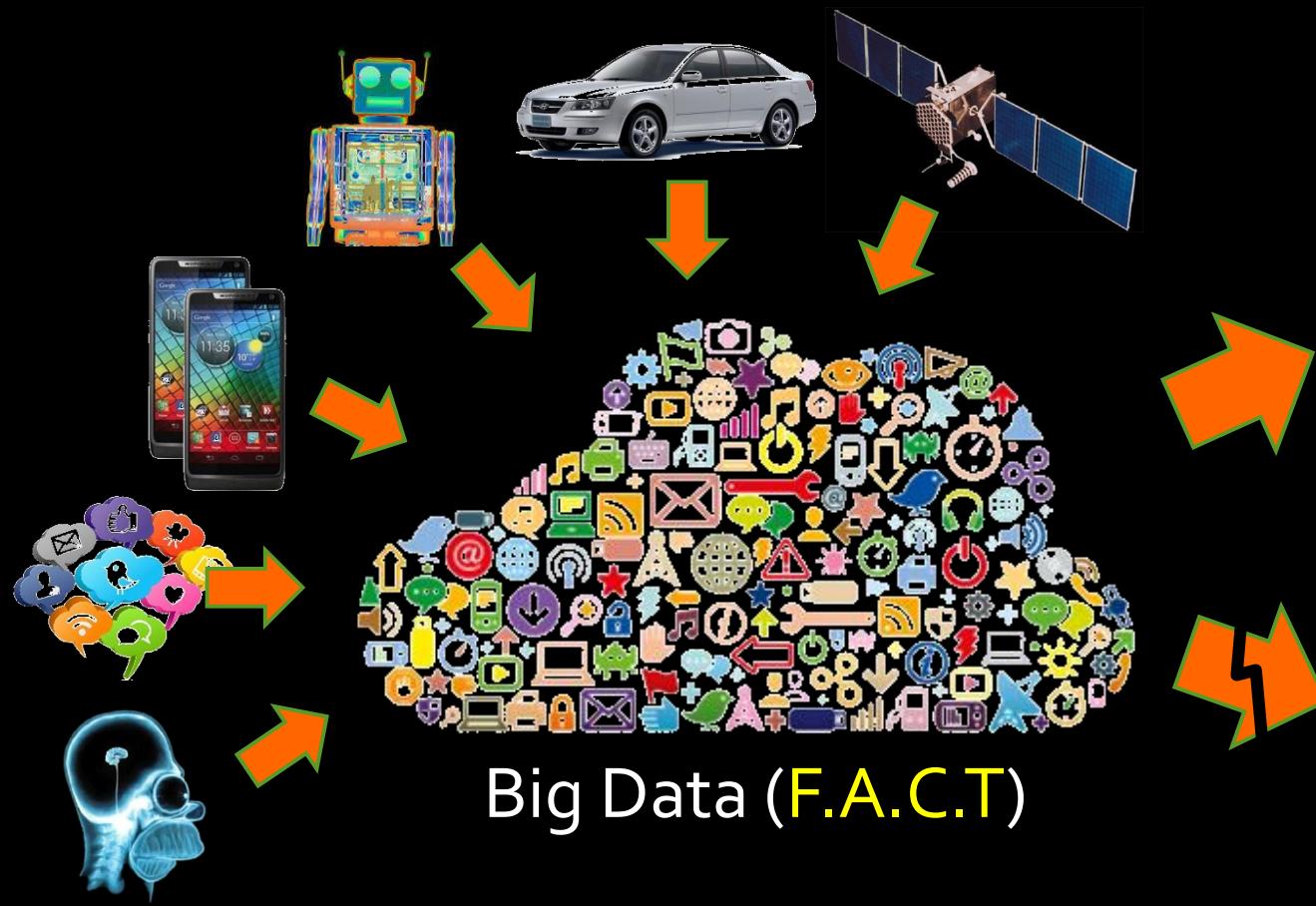
Pre-Historic Era
(12,000BC~3,000BC)



Historic Era
(~1,900AD)

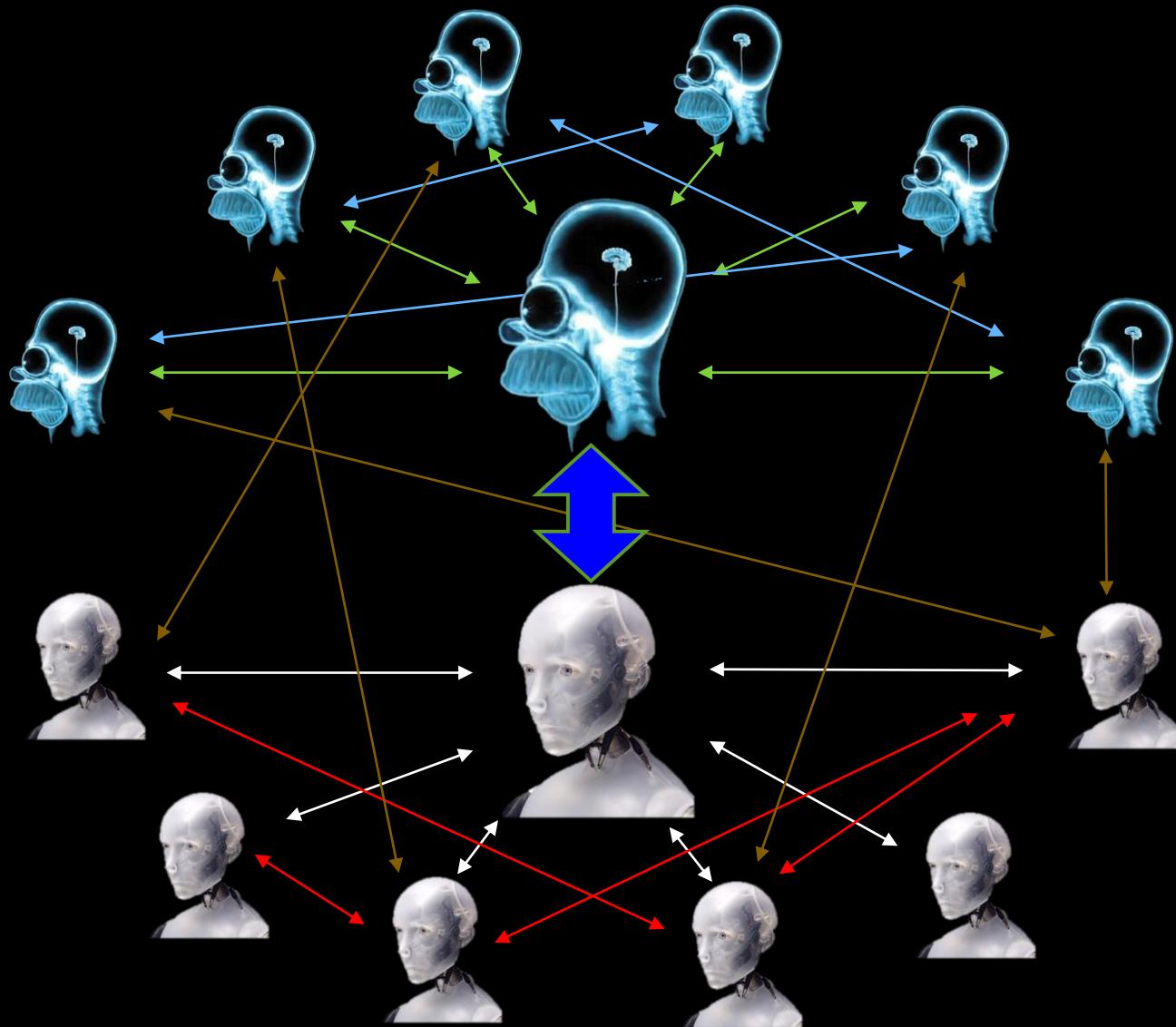


Big Data Era (2000~)



Augmented Brain?

New Collective Intelligence Era (2020~)



Hyper Connected People

Hyper Connected Devices

Hyper Connected Machines

Hyper Connected Services

Hyper Connected Data

Hyper Connected Knowledge

Act Two

Artificial Intelligence

God created 'Human being' based on their characteristics.

Human have a anxiety to create a 'Machine' like us?

A Space Odyssey (1968/2001)



Matrix (1999/2199)



I-Robot (2004/2035)



EX Machina (2015/2020?)



Terminator (1984/2029)

AI (2001/2090)

HER (2013/2025)

Don't let AI take our jobs (or kill us): Stephen Hawking and Elon Musk sign open letter warning of

- Letter says there is a 'broad range of potential risks'
- Areas benefiting from AI relate to health, transport and finance
- But in the short term, it will bring benefits
- In the long term, robots could become more intelligent
- Elon Musk has previously said AI could be like summoning the devil

By ELLIE ZOLFAGHARIFARD FOR DAILY MAIL

PUBLISHED: 19:08 GMT, 12 January 2015 | UPDATED:

 Share  

Artificial Intelligence has been described as the 'greatest threat to our existence'. Now a group of scientists and entrepreneurs have signed an open letter promising to ensure AI is used for the benefit of all.

The letter warns that without safeguards, AI could lead to a dark future.



Elon Musk worries AI could delete humanity along with space exploration

CEO of SpaceX and Tesla suggests AI might one day decide the solution for humans.

by Steven Musil  @stevenmusil / October 8, 2014

 0 /  3.6K /  508 /  8

Elon Musk has made no secret of his worries about the possible destructive power of artificial intelligence.

The billionaire chief executive of SpaceX and Tesla Motors may be a techno-optimist when it comes to solar power, space exploration and electric cars, but he continues to express his concerns that superintelligent machines might one day pose a threat to human existence.



Tesla

Microsoft's Bill Gates insists AI is a threat

By Kevin Rawlinson
BBC News



GETTY IMAGES

Bill Gates said he could not understand why people were not concerned by AI

Humans should be worried about the threat posed by artificial intelligence, Bill Gates has said.

The Microsoft founder said he didn't understand people who were not troubled by the possibility that AI could grow too strong for people to control.

Mr Gates contradicted one of Microsoft Research's chiefs, Eric Horvitz, who has said he "fundamentally" did not see AI as a threat.

Related Stories

[AI won't run amok, says Microsoft](#)

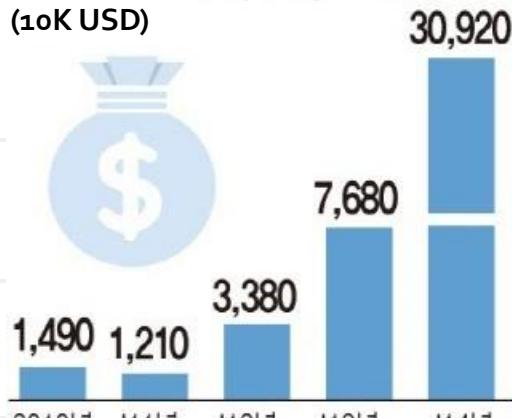
[Does AI really threaten the future of the human race?](#)

[Hawking: AI could end](#)

Artificial Intelligence Hypes?

VC investment in AI

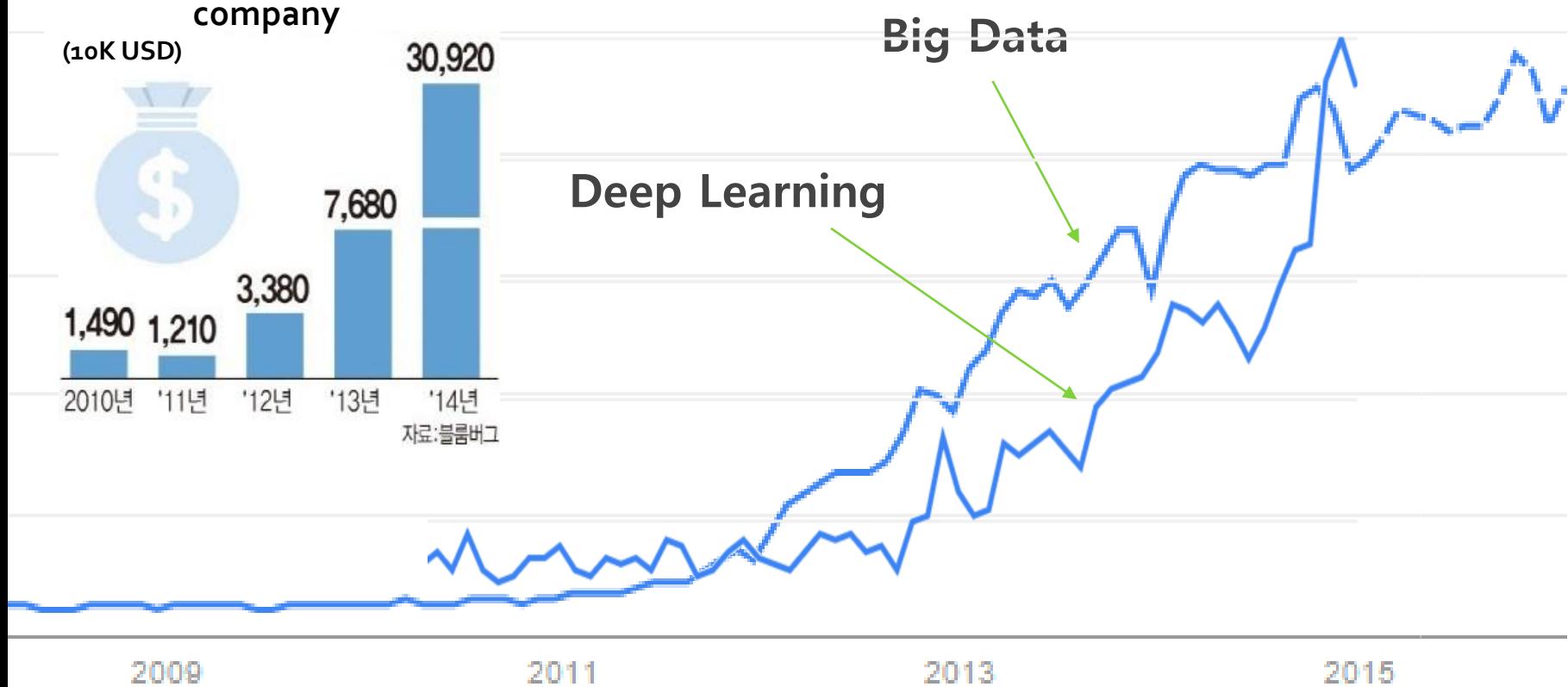
company
(10K USD)



Big Data

Deep Learning

자료:블룸버그



HISTORY OF AI & Big Data

350 B.C.
Aristoteles
Predicate Logic, Ontology 개념 제안

1629
Rene Descartes
공통이 기반 자동번역 아이디어

**Before
1930s**

150 B.C.
Antikythera mechanism
세계 최초의 아날로그 컴퓨터

1837
Charles Babbage
최초의 범용 컴퓨터
Analytical Engine 설계

1941
Alonzo Church
Lamda(람다) 계산 해설서 발행

1943
Warren McCulloch,
Walter Pitts
신경망(Neural Network)의
기초를 세움

1945
Alan Turing,
Alonzo Church
처치-튜링 명제
(Church-Turing Thesis) 발표

1949
Donald Hebb
뉴런간의 연결 정도를
변화시킬 수 있는 학습 규칙
(Hebbian Learning Rule)을 제안

1940s

1944
만하탄 프로젝트(美)
핵연쇄반응의 컴퓨터 시뮬레이션 수행

1949
정보이론
'정보의 아버지' C. Shannon
펀치카드, 사진 등의 대용량 데이터
저장 연구를 수행함

1950
Alan Turing
Computing Machinery and Intelligence 출간

Isaac Asimov
I Robot(나는 로봇) 출간

1951
Minsky, Edmond
SNARC(신경회로망 컴퓨터)

1952
Bell Labs
최초의 음성인식기 개발

1954
Yehoshua Bar-Hillel
최초의 자동번역 엔진 구현

1955
Allen Newell, J.C. Shaw,
Herbert A. Simon
최초의 AI 프로그램 Logic
Theorist (LT)를 선보임

1956
J. McCarthy, M. Minsky, A. Newell, H. Simon
Dartmouth workshop에서 인공지능(AI) 용어를 최초로 사용

1958
John McCarthy, Marvin Minsky
프로그램 언어인 Lisp 발명

1950s

1950
범용 컴퓨터 ENIAC
최초의 범용 컴퓨터 ENIAC을 날씨 예측에 활용함

1957
데이터스트리 알고리즘
'최단 경로 문제(SPP)' 해결 알고리즘 발표(Dijkstra)

1958
컴퓨터의 예측모델
FICO사(美)가 신용리스크 결정에 예측모델을 적용 시작



1960
Bernard Widrow, Marcian Hoff
퍼셉트론 모델의 선형화와 유사한
워드로 호프 모델을 제안

1962
Frank Rosenblatt
퍼셉트론 수렴이론
(Perceptron Convergence Theorem)을 발표

1963
M. Ross Quillian
지식표현 수단으로서의 Semantic Networks을 소개

Marvin Minsky
Steps Towards Artificial Intelligence 논문 발표

1965
Bruce Buchanan,
Edward Feigenbaum,
Lederberg
DENDRAL 전문가시스템 프로젝트

MIT ELIZA
최초의 자연언어처리 시스템

1960s

1962
IBM 1311 HDD
작동 가능한 최초의 상용 하드디스크
세탁기 크기로 2MB이트 저장 가능

1964
CODASYL / IDS
최초의 상용 DBMS

1967
데이터 압축 알고리즘
BA Marron이 '자동 데이터 압축'에
관한 논문을 ACM에 발표함

Automatic
Data
Compression

1970
Jaime Carbonell
SCHOLAR

1972
Alain Colmerauer
컴퓨터 언어

1973
Edward H. Adelson
감염성 혈액
항생제를 찾는

1974
Paul Werthimer
역전파 (Backpropagation)
신경망의 기계학적

1976
MRP 시스템
Material Requirements Planning
컴퓨터가 본

1970

1970
Edgar F. Codd
ROB 모델 제안

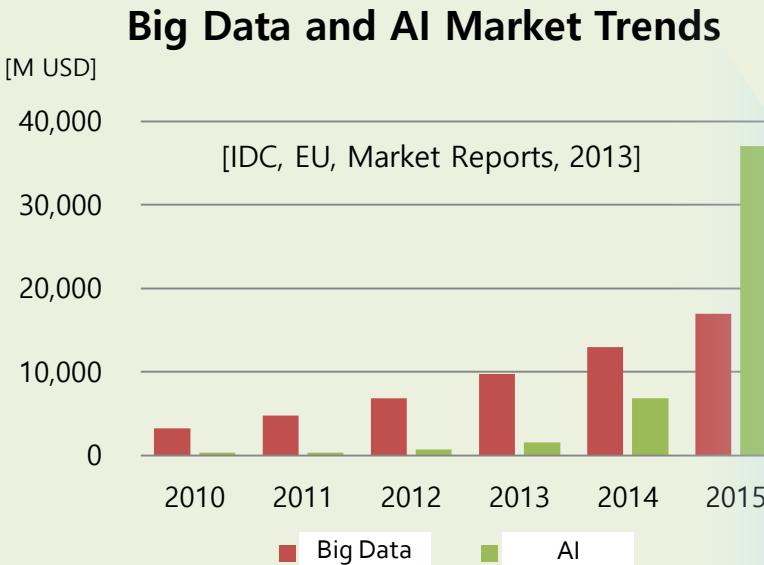
1974
SQL
IBM에서 최초로 도입

1977
통계컴퓨팅
국제 통계학회
애플 개인용 컴퓨터
최초의 대량화된 (1MHz CPU)

1978
Oracle 출시
Larry Ellison

1979
관계형 DBMS
최초의 상용화

Big Data and AI Market



- Big Data market will grow at a 27% compound annual growth rate (CAGR) to \$32.4 billion through 2017
- Global Industry Analysts forecast the big data based Artificial Intelligence market to exceed €27 billion by 2015

Application Markets



Smart Robot / Wearable Services



Virtual Agent / Decision Support



Self-driving Car / Smart Factory

Machine Intelligence LANDSCAPE

CORE TECHNOLOGIES

ARTIFICIAL INTELLIGENCE



DEEP LEARNING



MACHINE LEARNING



NLP PLATFORMS



PREDICTIVE APIs



IMAGE RECOGNITION



SPEECH RECOGNITION



RETHINKING ENTERPRISE

SALES



SECURITY / AUTHENTICATION



FRAUD DETECTION



HR / RECRUITING



MARKETING



PERSONAL ASSISTANT



INTELLIGENCE TOOLS



RETHINKING INDUSTRIES

ADTECH



AGRICULTURE



EDUCATION



FINANCE



LEGAL



MANUFACTURING



MEDICAL



OIL AND GAS



MEDIA / CONTENT



CONSUMER FINANCE



PHILANTHROPIES



AUTOMOTIVE



DIAGNOSTICS



RETAIL



RETHINKING HUMANS / HCI

AUGMENTED REALITY



GESTURAL COMPUTING



ROBOTICS



EMOTIONALrecognition



HARDWARE



DATA PREP



DATA COLLECTION



SUPPORTING TECHNOLOGIES

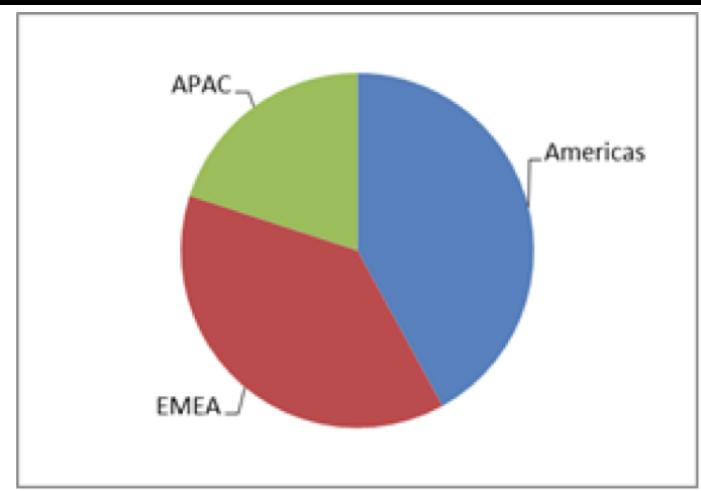
AI Market Expectation?

- Growing exponentially over the coming years
- Apple, Google, Microsoft and Facebook invest in the AI



Virtual Agent Market

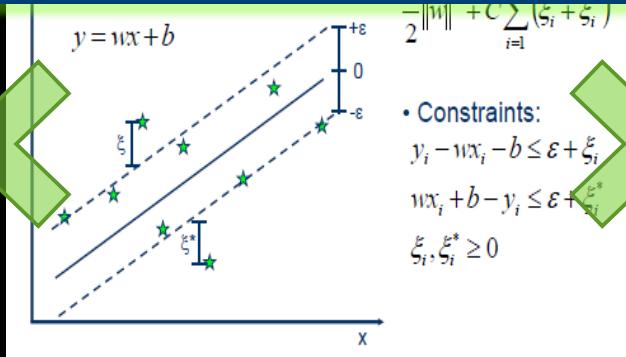
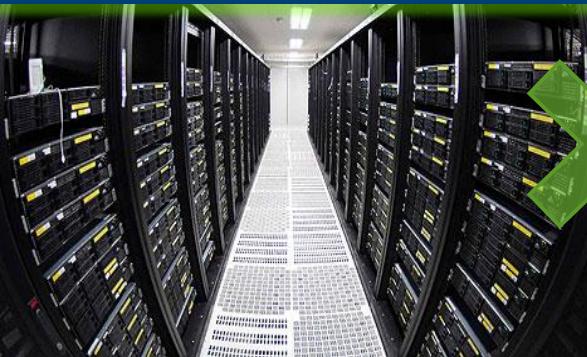
- BusinessWire (2015)
 - Global Market size is growing 30% per year: 0.35 (2012) to 3.1 billion dollars (2020)
 - Communication service to O2O(Online 2 Offline) service with Q&A
 - Korean market will grow up to 0.24 billion dollars (2020)
- GrandViewResearch (2015), TechNavio (2014)
 - North America centric → Europe and Asia market is growing rapidly, up to 20% in 2018



* TechNavio, 2014



Going to be almost FREE! (except data...)



**big
data**

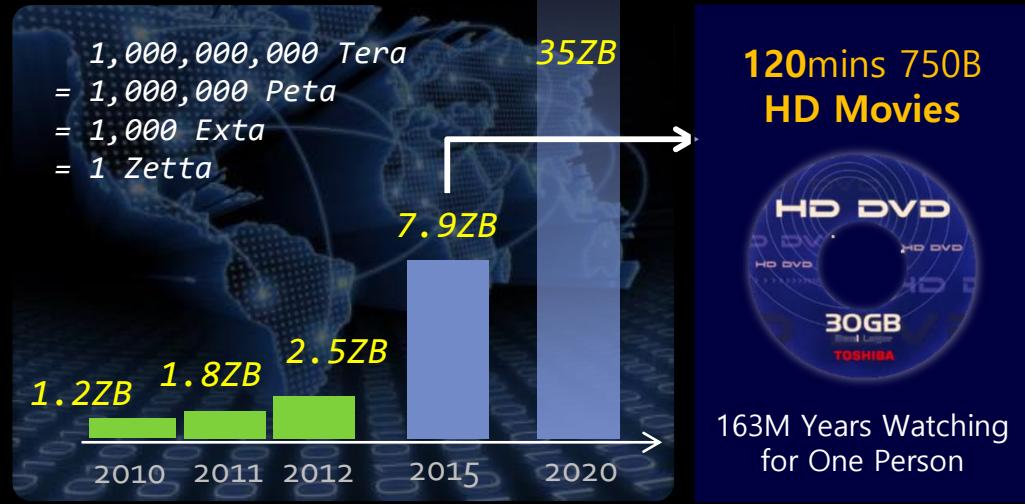
Cloud Infrastructure

AI Algorithms

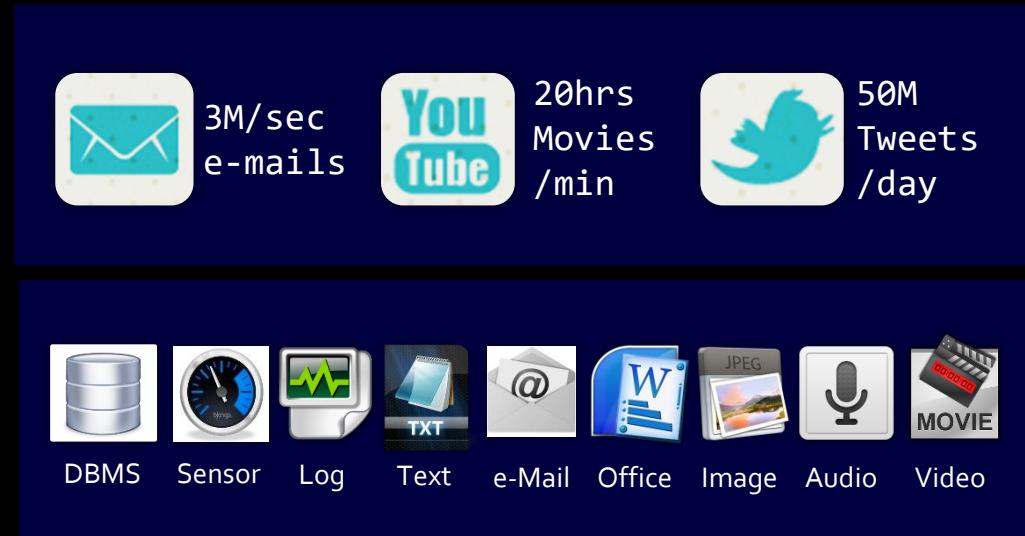
Big Data

Big Value from Big Data?

Size?



3V? { Volume Velocity Variety



The **Facts** of Big Data?

Big Data Technology :

“Complex and large data sets that it becomes difficult to process using traditional technologies”

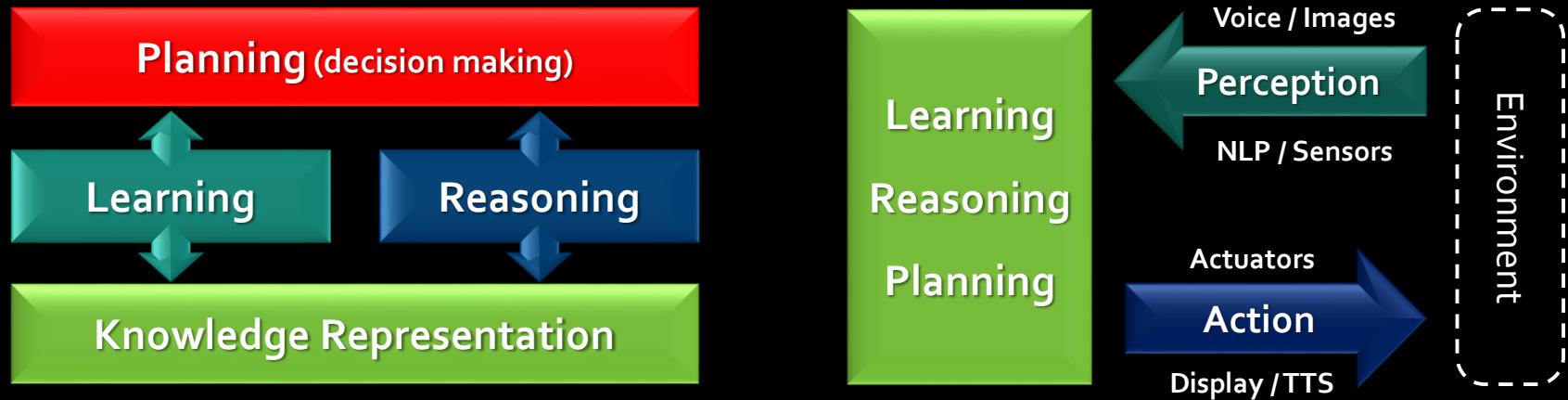
Why is it so difficult to process?

3V x F.A.C.T !!

(Fragment x Ambiguity x Context x Trustability)

What is Artificial Intelligence?

- The science and engineering of **making intelligent machines**.
(John McCarthy)
- The study and design of **intelligent agent system** that perceives its environment and takes actions that maximize its chances of success. (Stuart Russell)
- The study how to create computers and computer software that are **capable of intelligent behavior**. (Wikipedia)



What is Artificial Intelligence?

- **Strong AI**

Artificial general intelligence. Computers can be made to think on a level at least equal to humans, that they can be **conscious** and experience emotions.

- **Weak AI**

Non-sentient computer intelligence or AI that is **focused on one narrow task**. All real-world systems labeled "artificial intelligence" of any sort are weak AI at most.

Systems that think like humans	Systems that think rationally
Systems that act like humans	Systems that act rationally

Knowledge Representations?



Knowledge Representations

Natural Language

"Employees working for a company are humans; the company and the employees are legal entities. The company is able to make a reservation for an employee's trip. The trip is available by plane or train that travels in cities within Korea or the U.S.. The companies and destinations for business trip are located in the cities. Saltlux reserved OZ510 with a round trip of Seoul and New York for Hong, Kildong."

Rule Language

(Rule) If someone is flying, he must be on trip.

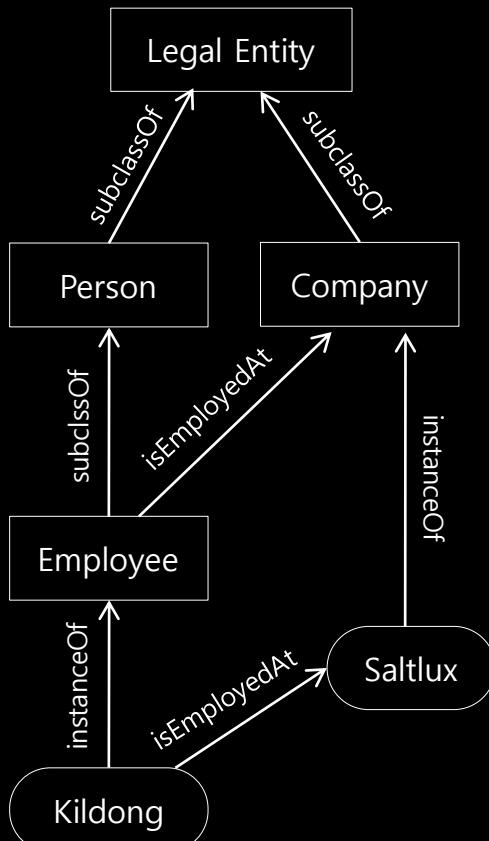
(Rule) If someone's trip is reserved in a company, he is an employee of the company.

(+ Rule) For short trip in the same country, an employee should take a train.

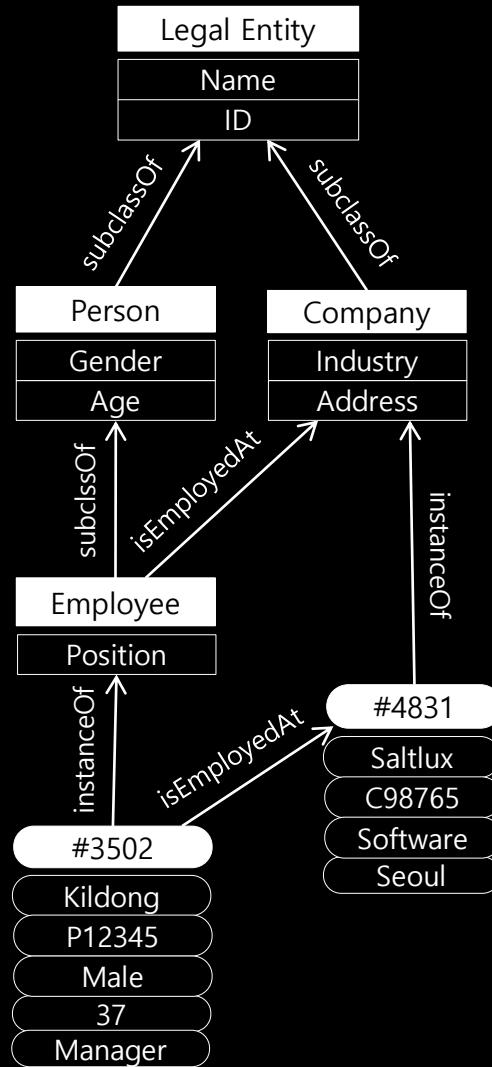
(Deduction) Hong kil-dong whose flight is in reservation is an employee of Saltlux.

(Deduction) OZ510 is a flight for the U.S. and Korea.

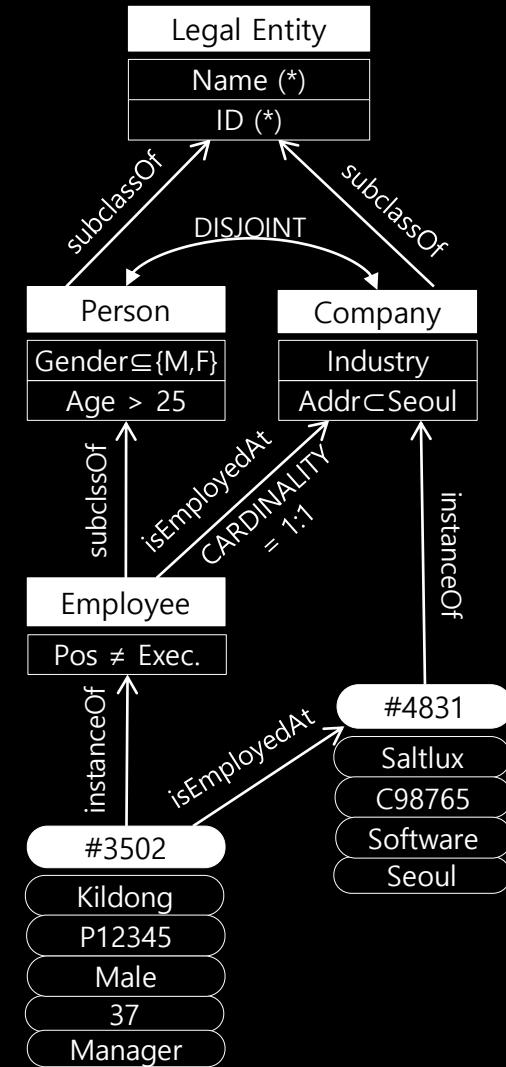
Building Explicit Knowledge Base



(a) Semantic Network

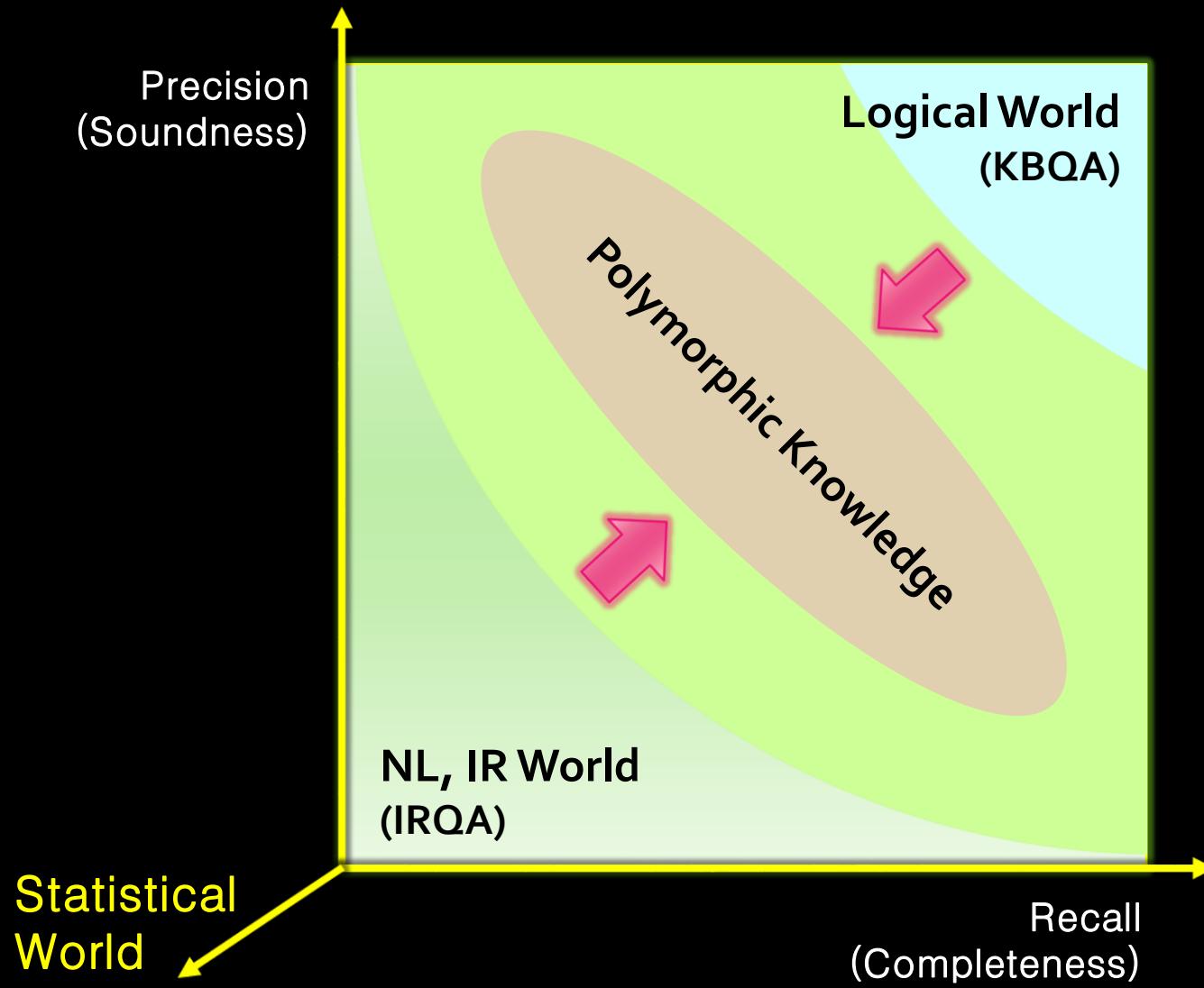


(b) (a) + Frame (Slots)



(c) (b) + Logical Restrictions

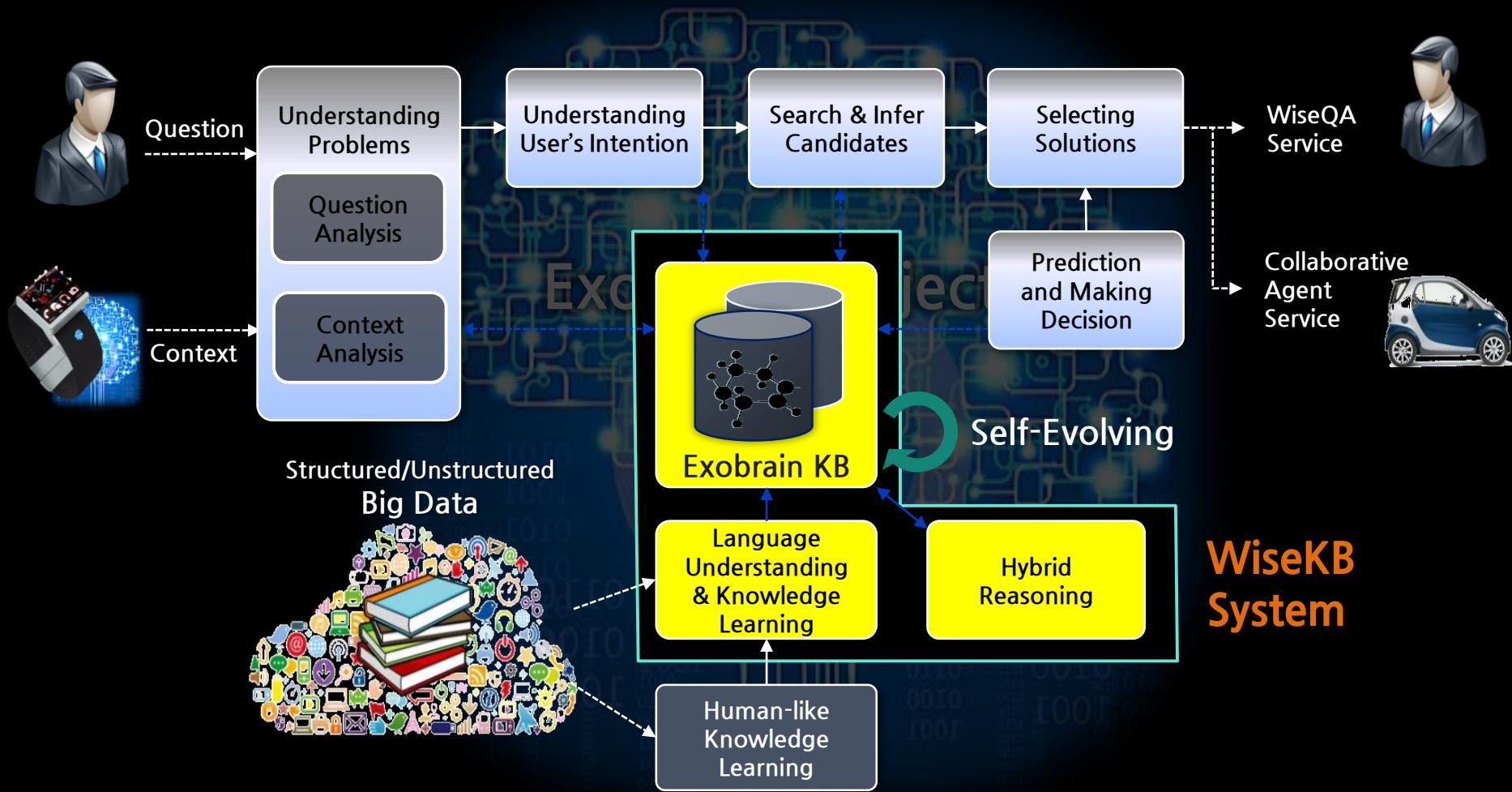
Building Polymorphic Knowledge Base



Act Three

Exobrain and WiseKB

Exobrain is a future AI that can learn, think and make decision like human



Knowledge Learning and Reasoning for QA



Who is this person? He/she was born in Gyungsang-do between remarried father, a soldier and mother, a teacher. He/she has an elder sister and lost his/her mother on independence day.

WiseQA

Whols(?x) :-
 hasMother(?x, ?mother),
 job(?mother, teacher),
 becomes(?x, 60),
 home(?x, Gyungsang-do),
 hasFPosition(?x, 2nd),
 hasFather(?x, ?father),
 job(?father, soldier),
 remarriage(?father, ?mother)

Ara Ko,
 Jeaho Kim,
 Jun Han...

Search
 (graph-matching)

G.H. Park(90%)
 Ara Ko(10%)
 Jaeho Kim(5%)
 ...

Reasoning
 Semantic
 Temporal
 Geospatial
 Uncertain

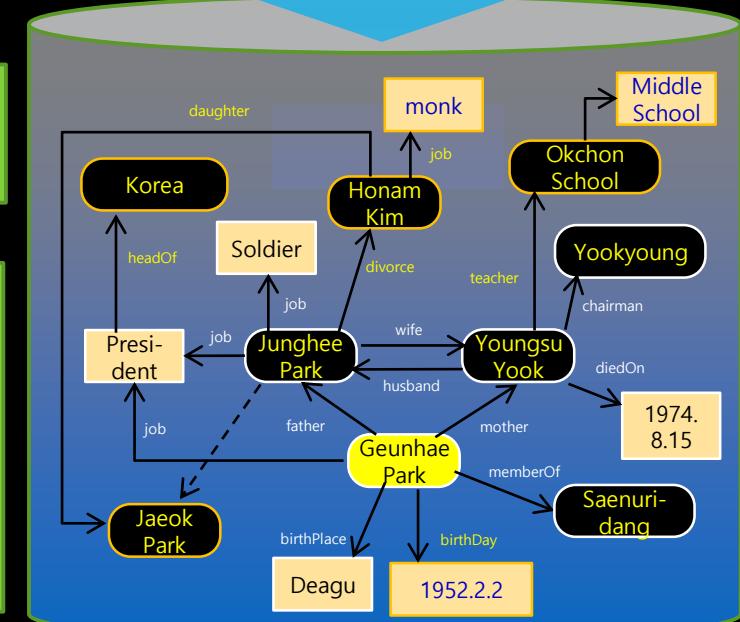


박근혜는 1952년 2월 2일 경상북도 대구시 삼덕동(현재의 대구광역시 중구 삼덕동 5-1번지, 5-2번지 일대)에서 전시(戰時) 대구 주재 육군본부 작전·교육국 작전차장 박정희 대령과 중등학교 교사 출신인 그의 부인 육영수의 딸로 태어났다. 어머니 육영수에게는 첫 소생이었으나 아버지 박정희는 이혼 경력과 전처소생의 장녀 박재옥이 있었으므로 박정희에게는 차녀가 된다. 여동생 박근령과 남동생 박지만이 있다.

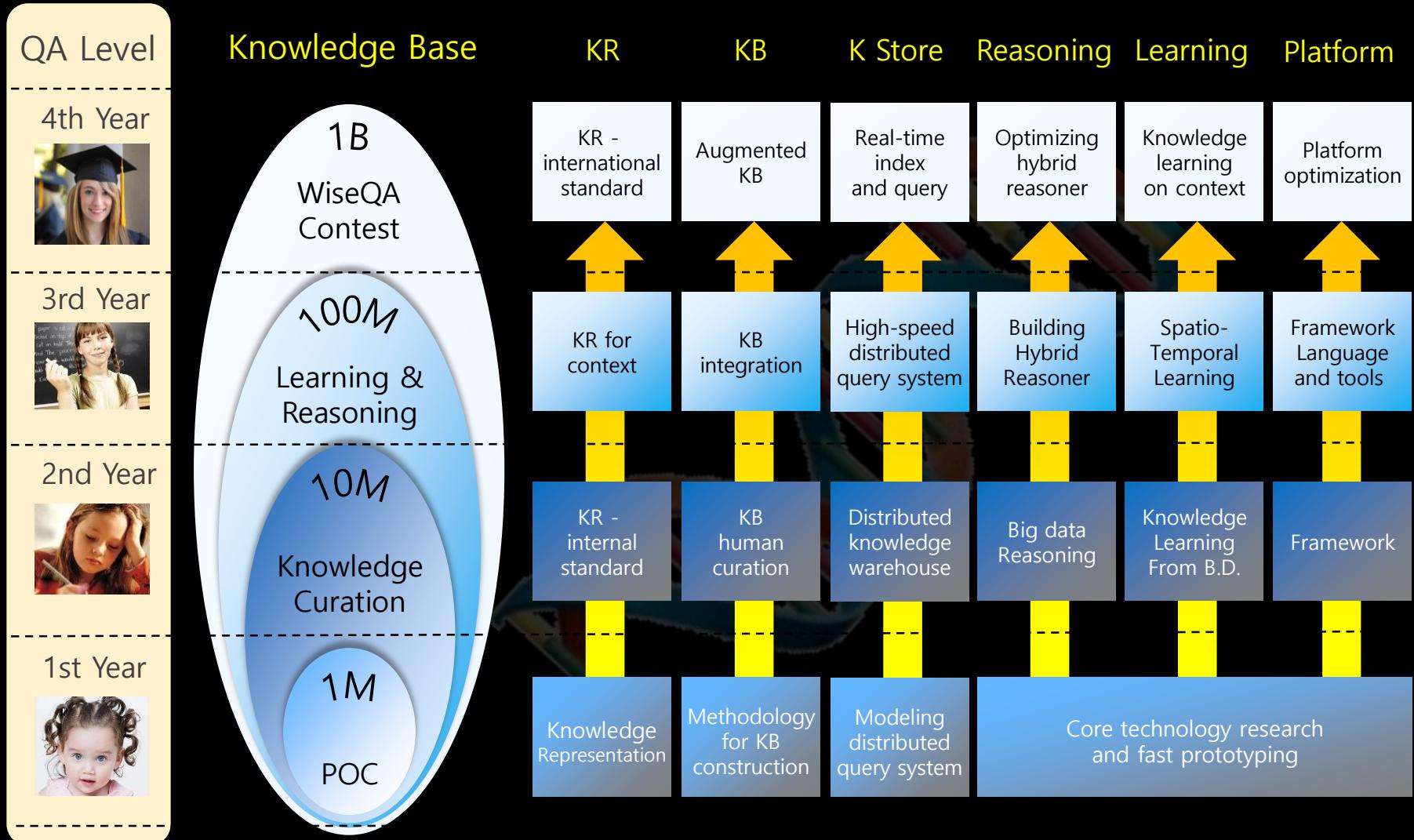
국적 : 대한민국
 출생일 : 1952년 2월 2일 (61세)

위대한 시도자 박정희 장녀 박근혜님 대통령 당선!!!

Learning and
 Augmentation



R&D Milestones and Goals



WiseKB - Distinguished Features



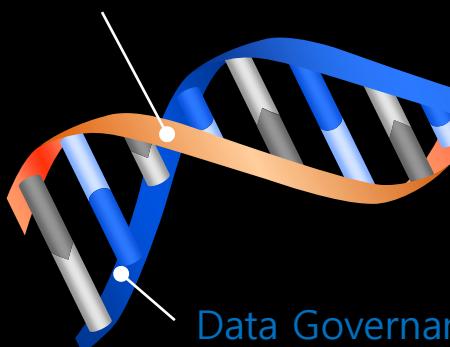
Feature	Feature	Comparison	WiseKB Uniqueness
P	Knowledge Representation	<ul style="list-style-type: none">• CyCorp Cyc• IBM Watson	<ul style="list-style-type: none">• Integration of logical, linguistic and statistic knowledge
	High Quality and Volume	<ul style="list-style-type: none">• Google Graph• Wolfram Alpha	<ul style="list-style-type: none">• Dual Spiral methodology for knowledge acquisition
S	Knowledge Learning	<ul style="list-style-type: none">• IBM Watson• Google Graph	<ul style="list-style-type: none">• Hybrid Learning (ML+Rules)
	Self-Verify and Proof	<ul style="list-style-type: none">• CMU NELL• Google Graph	<ul style="list-style-type: none">• Big data based confidence prediction
H	Diversity of Reasoning	<ul style="list-style-type: none">• EU LarkC• CyCorp Cyc	<ul style="list-style-type: none">• Semantic, geospatial, temporal and uncertainty reasoning
	Economic efficiency	<ul style="list-style-type: none">• VU WebPie• Franz AllegroG	<ul style="list-style-type: none">• Parallelism in memory, GPGPU

Dual-Spiral methodology for building KB



Knowledge Curation

- High quality KB
- Semantic annotation
- KB integration



Data Governance

- Acquisition of big data and knowledge resources
- Automatic resource discovering for lack-knowledge
- Semantic data integration

Crowd Sourcing

- Gamification (Quiz game)
- Acquisition and proving

Domain Experts

- QA tests
- Verifying learned KB

Human Centric ▶

Lifecycle Management for
Knowledge Augmentation

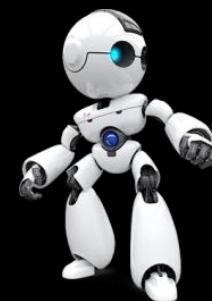
Machine Centric ▼

Self-Learning

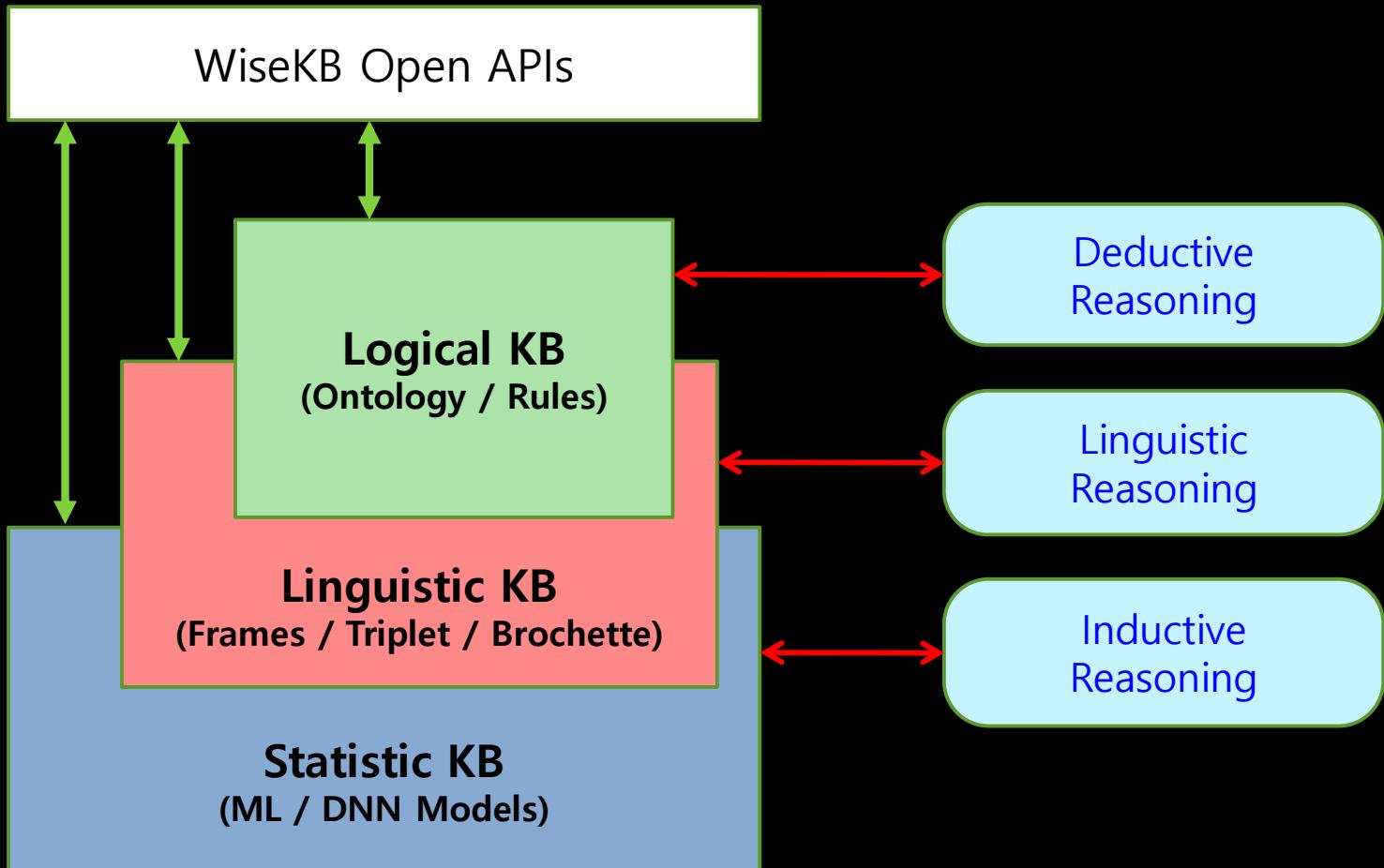
- Learning by reading
- Learning by taking advice

Reasoning

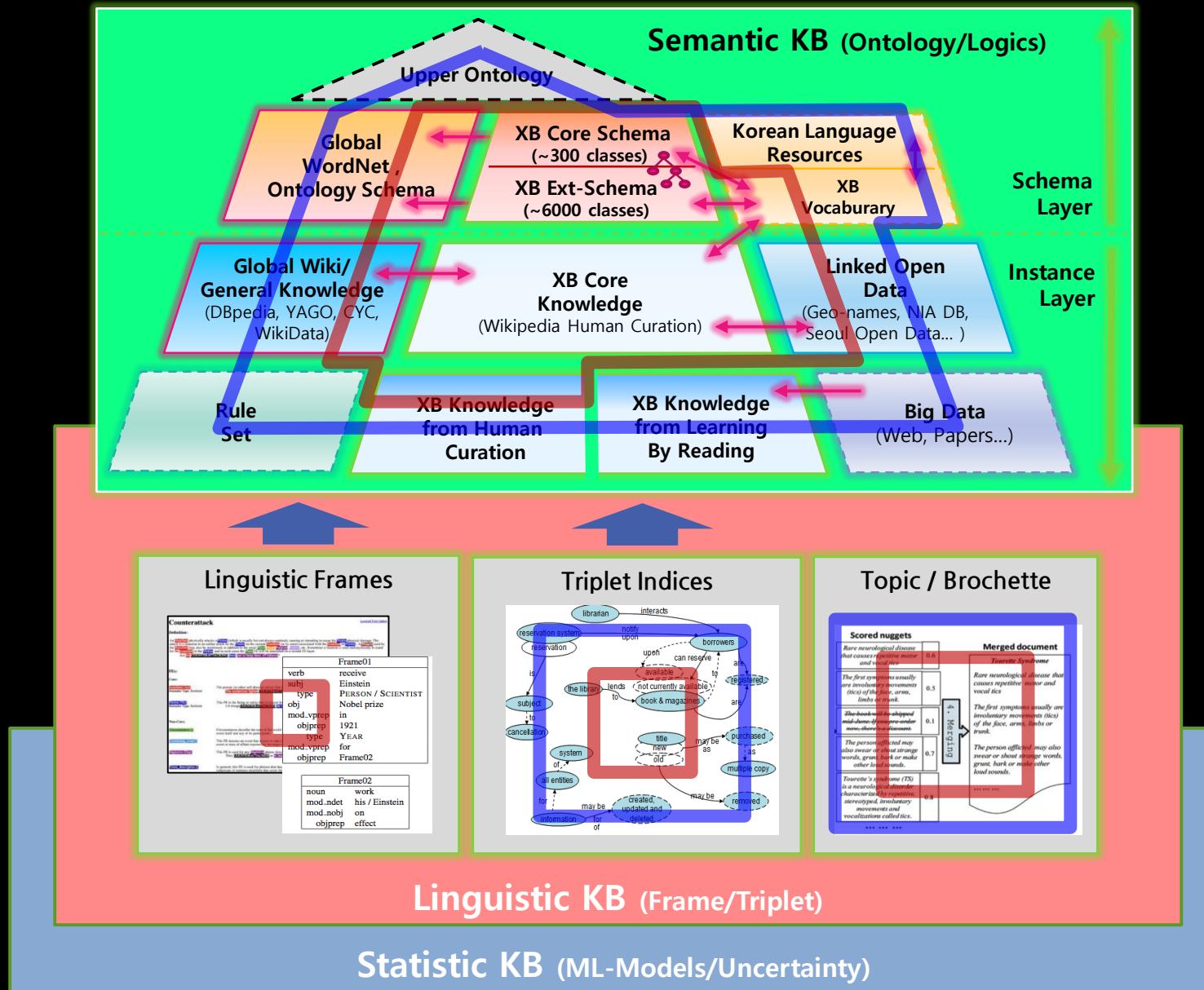
- Hybrid logic reasoning
- Spatio-temporal reasoning
- Statistic and uncertainty



Polymorphic Knowledge Base

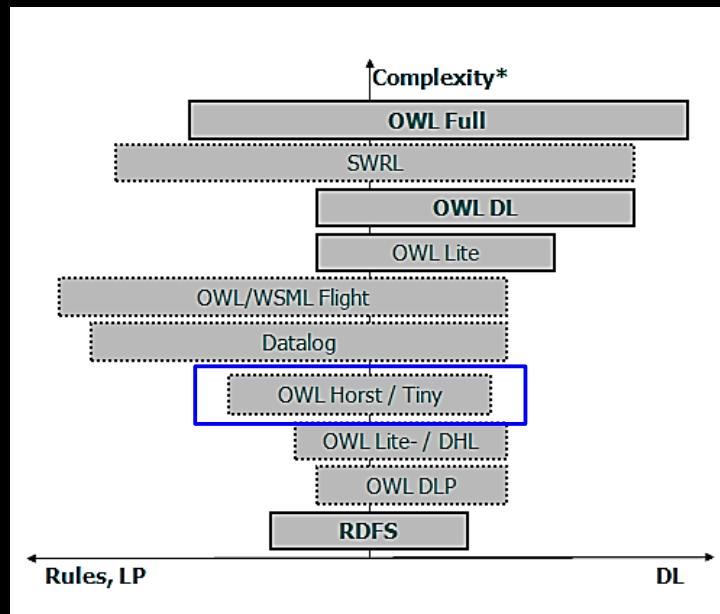


Exobrain KB for Polymorphic Knowledge and Hybrid Reasoning

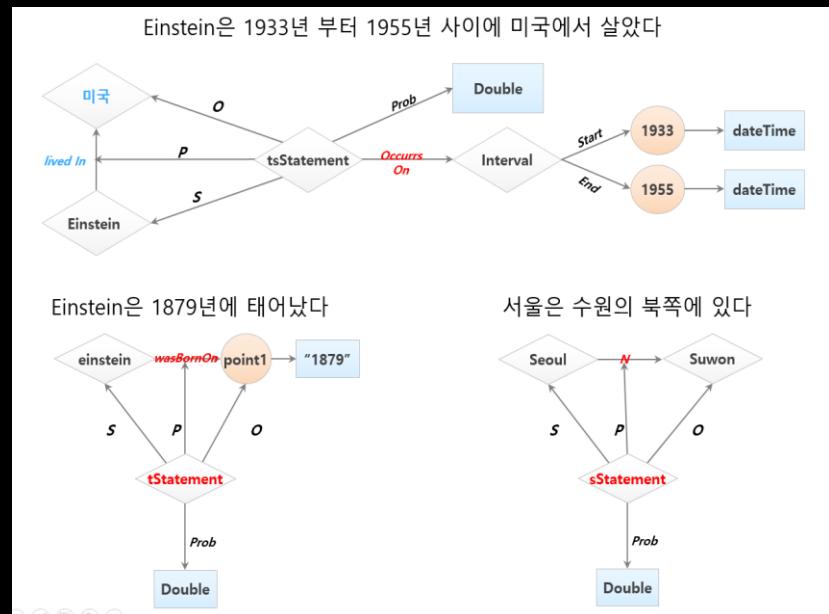


Logical Knowledge Representation

- Labeled graph/Triple(s p o) based KR
- Using reification for temporal, spatial and probabilistic knowledge
- Expressivity of OWL horst level
- SPARQL query language for KB accessing



Logical expressivity for KR



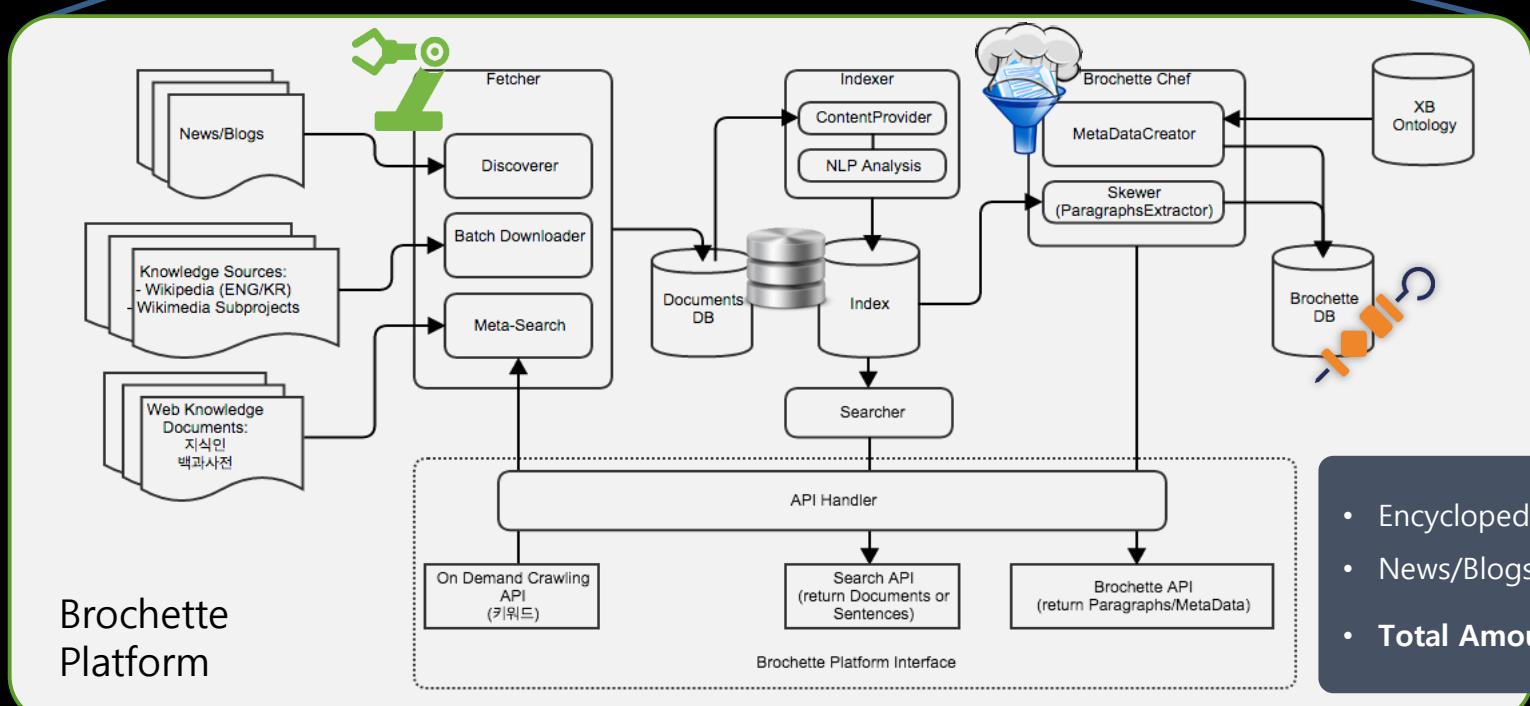
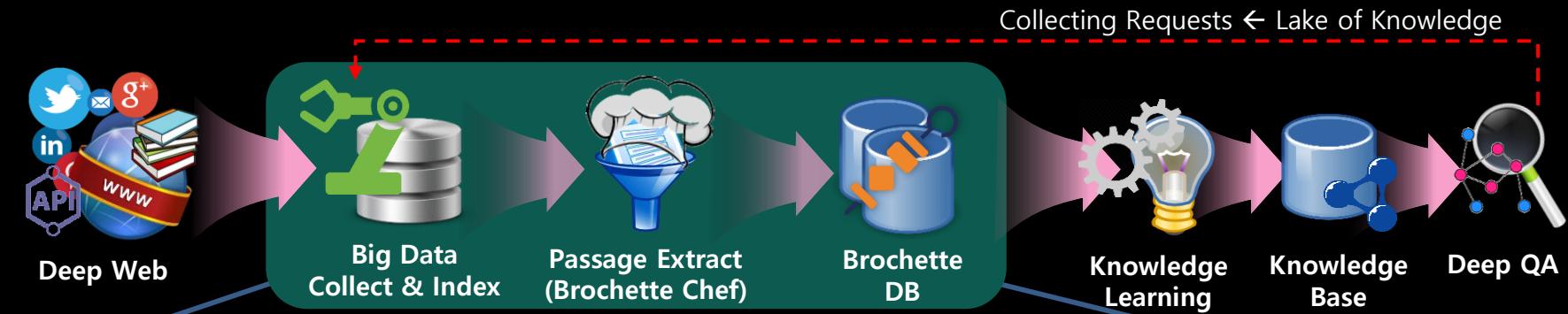
Reification for KR

Constructed XB - Core KB

- High quality and huge-scale knowledge base development based on dual spiral methodology
- Manual knowledge curation, semi-automatic knowledge importing and validation from Wikipedia and Linked Open Data.
- Current volume: 186M triples (Biggest KB in Asia)

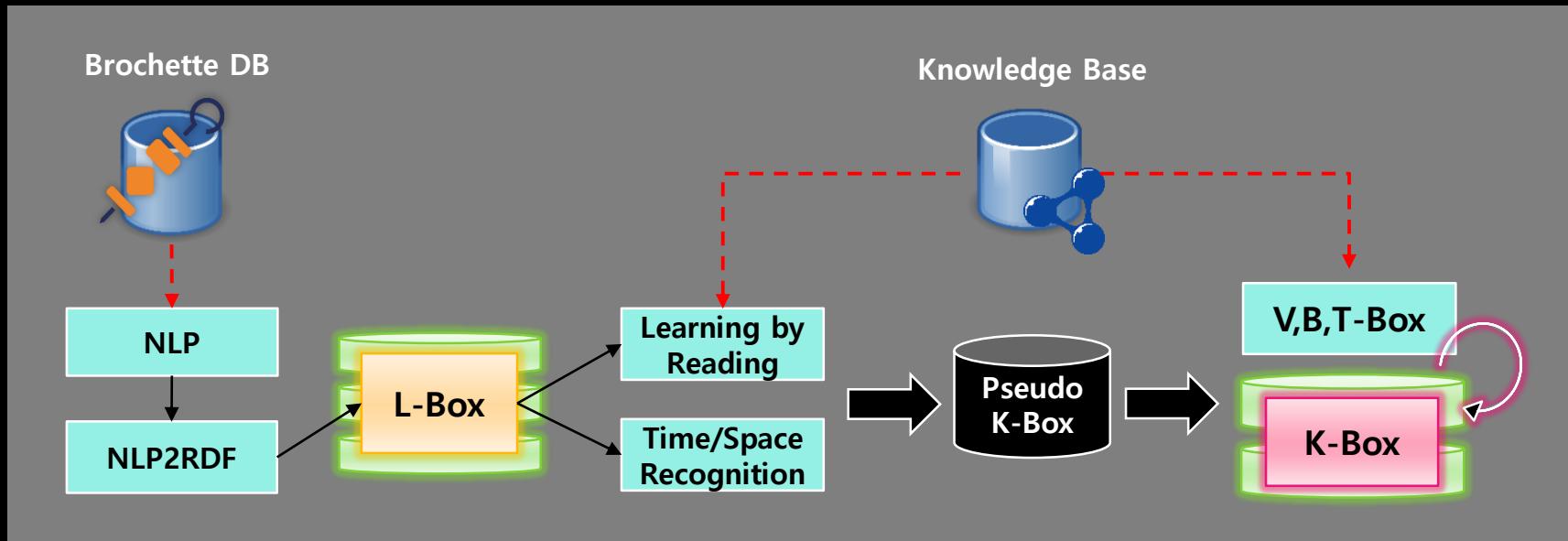
Type	2nd year	3rd year	Difference
Classe	6,132	6,315	▲183
Property	504	991	▲487
Instance	1,554,489	23,399,338	▲ 20M
Triple	10,639,996	186,000,000	▲80M
Domain	General(wiki), History, Person, Organization	+ General(news, blogs), GeoSpatial, Art work, Science, Events, and etc.	

Knowledge Resource Acquisition



- Encyclopedia: 2M articles
- News/Blogs: 72M articles
- **Total Amount: 370K Books**

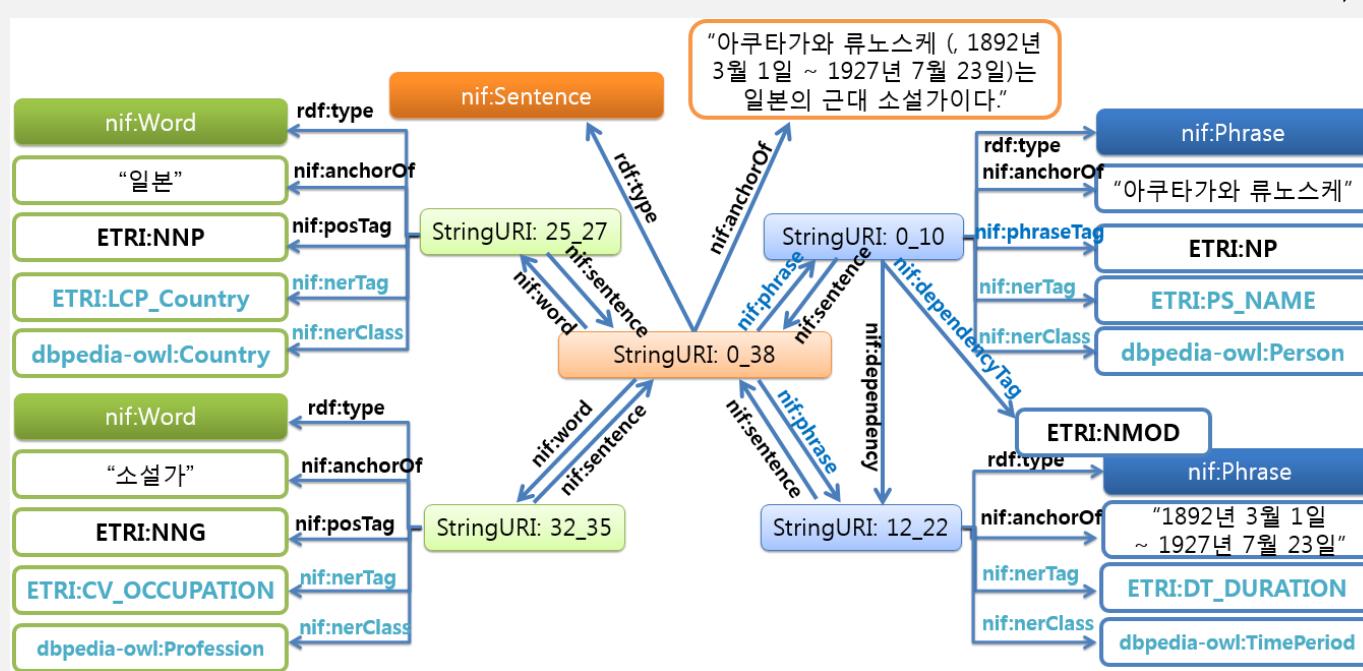
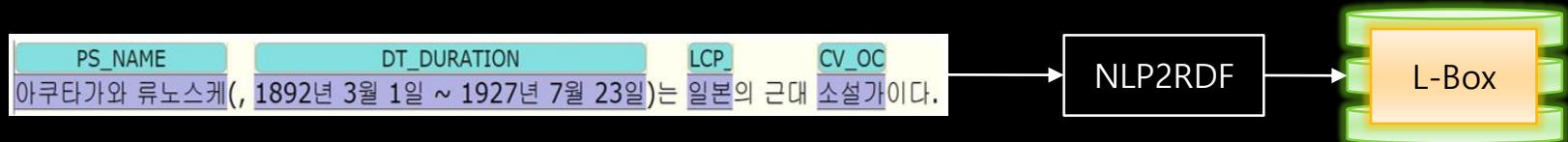
Knowledge Learning from Big Data



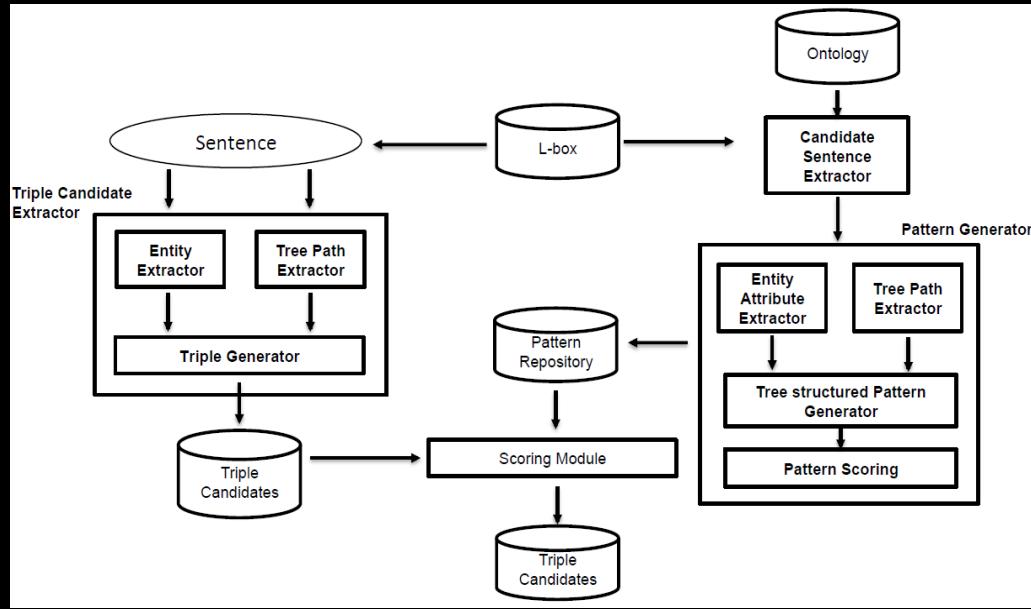
- Development of knowledge learning and automatic validation from unstructured big data
- Learning by Reading based on NLP and Machine Learning
- 1B triples in L-Box, 40M triples in K-Box with 79% accuracy of entity learning and 60% accuracy of relationship learning

L-Box construction for Knowledge Learning

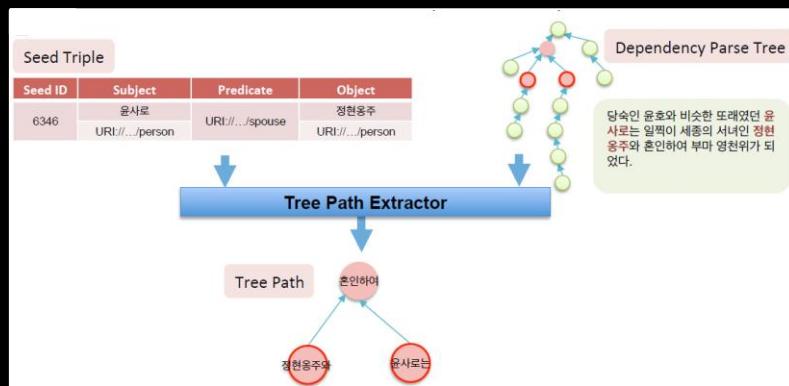
Integrate all NLP tools (morph, parser, NE ...) and convert results into unified RDF based NIF



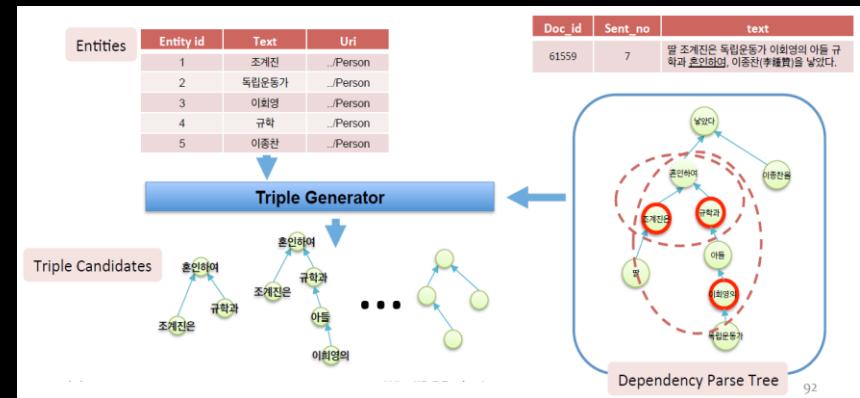
ML based Learning by Reading



- Learning knowledge patterns from seed triple and L-Box (70.8 % accuracy)
- Learning new knowledge (triples) from learned patterns and L-Box
- Knowledge learning from the body text of Korean Wikipedia (57.1% accuracy, 76% w/ p-error)



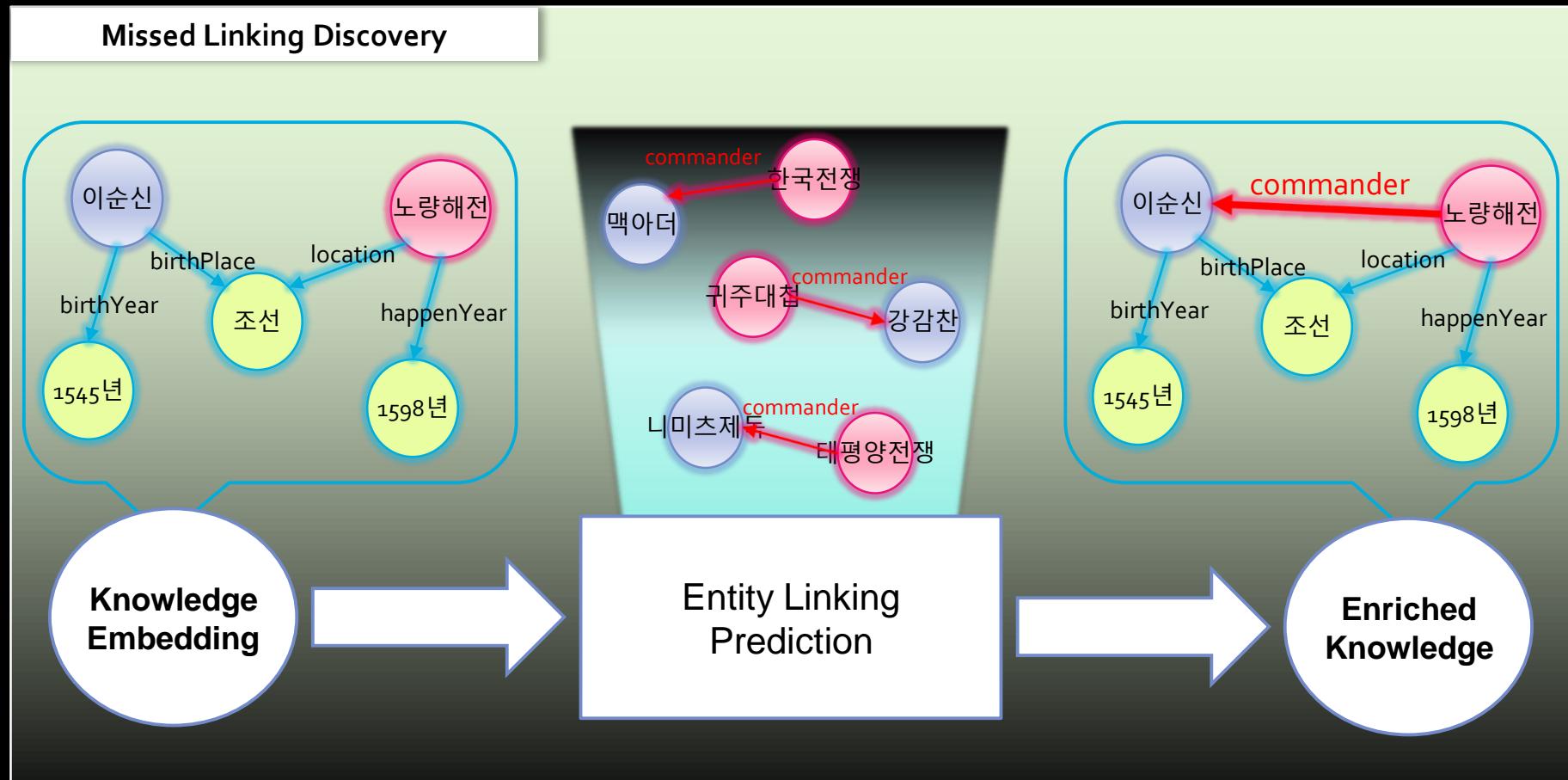
1. Pattern Learning



2. Knowledge Learning

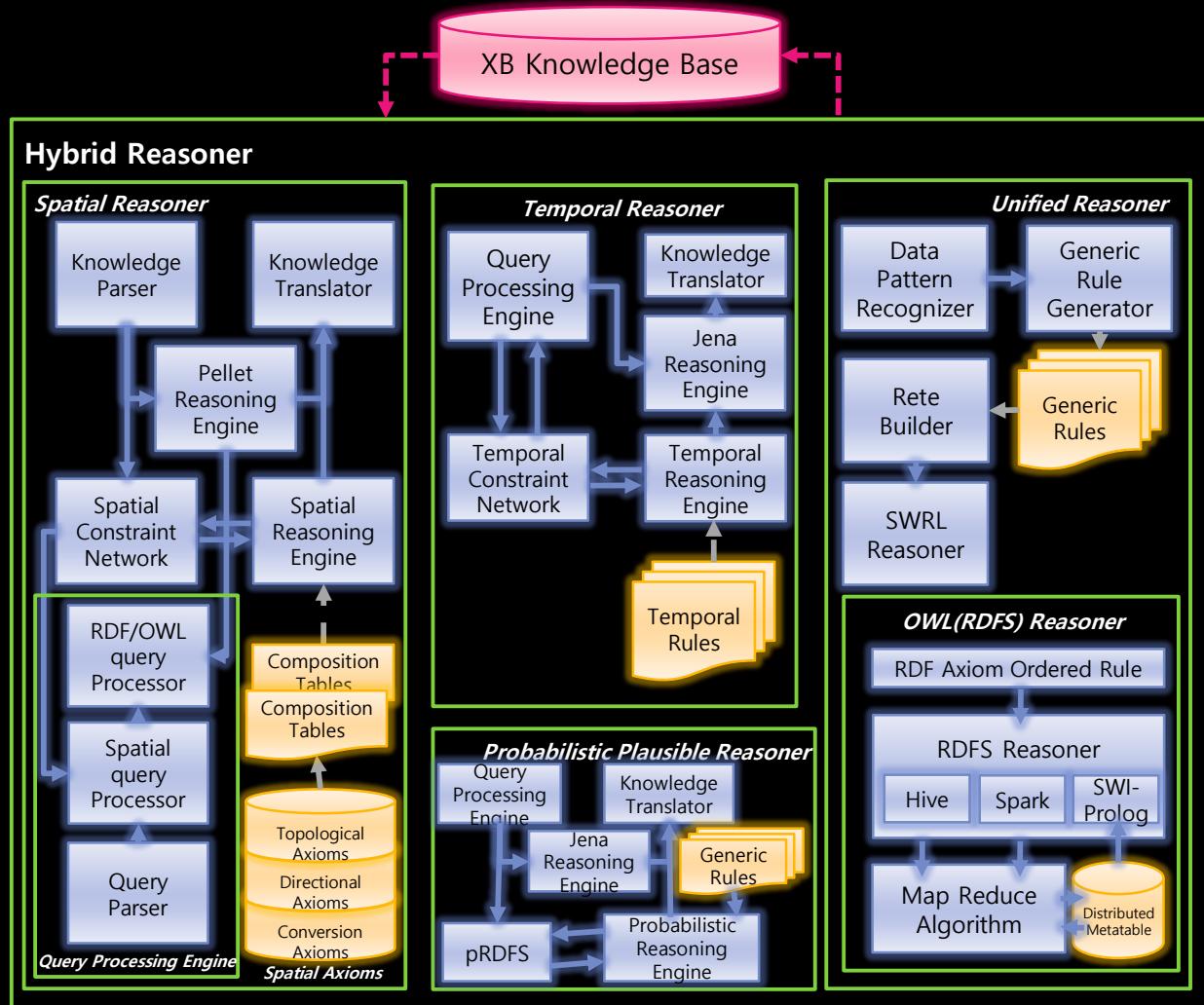
Knowledge Embedding and Link Prediction

Applying deep learning and knowledge embedding technology for missing link prediction (80% accuracy)



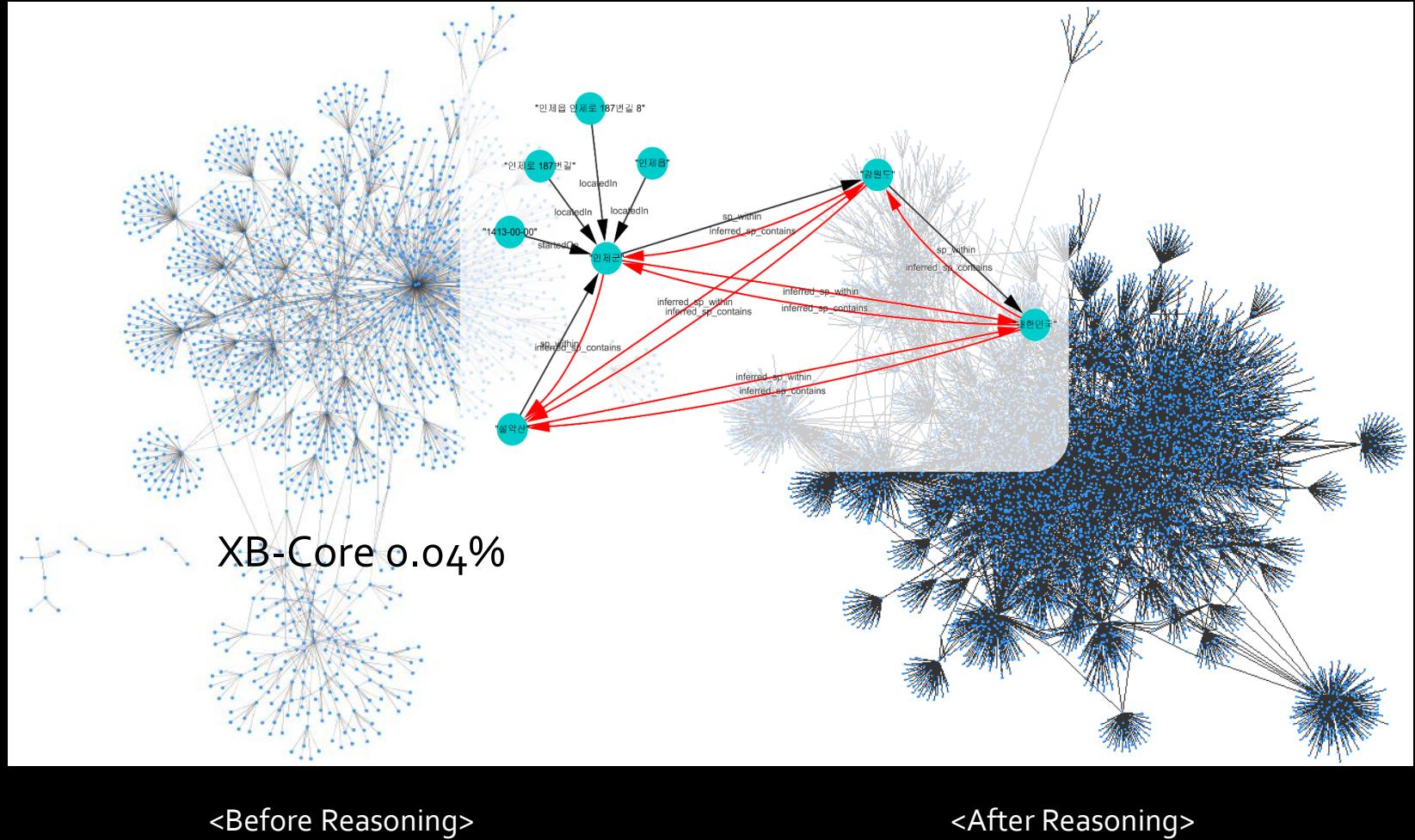
Hybrid Reasoning

- Huge scale and fast DL-horst reasoner working on in-memory and map-reduced architecture
→ World fastest reasoner
- Developing spatial and temporal reasoner based on CSD-9, RCC-8 and Allen's algebra
- 2.6M inferred triples from 3.4M XB ontology
- 460 k triples/s throughput on Spark



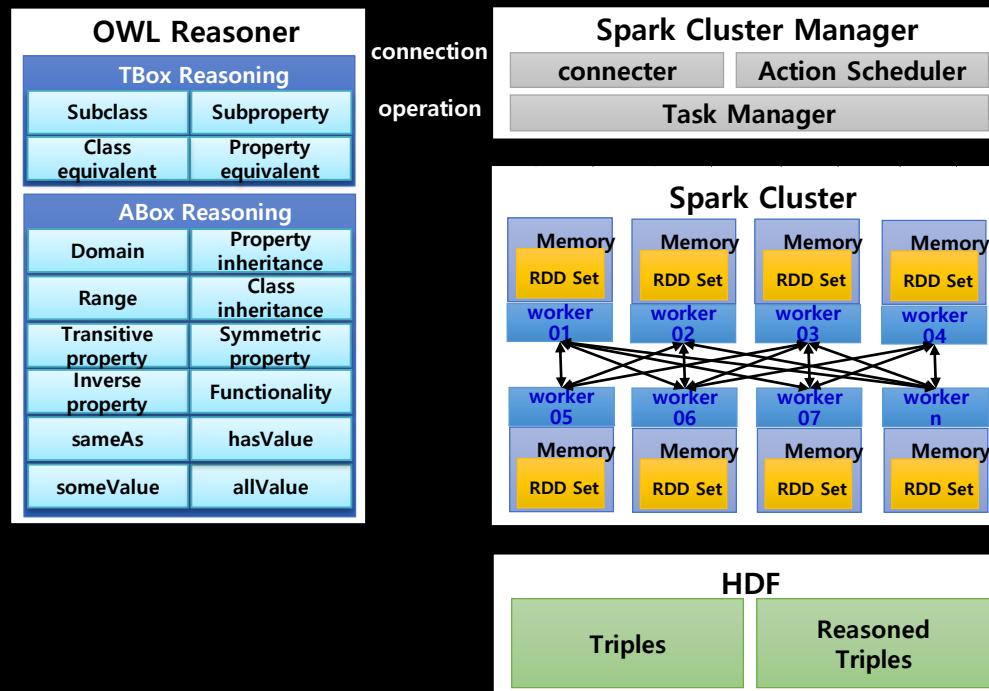
Semantic Reasoning and Knowledge Enrichment

World fastest reasoning performance based on OWL Horst expressivity

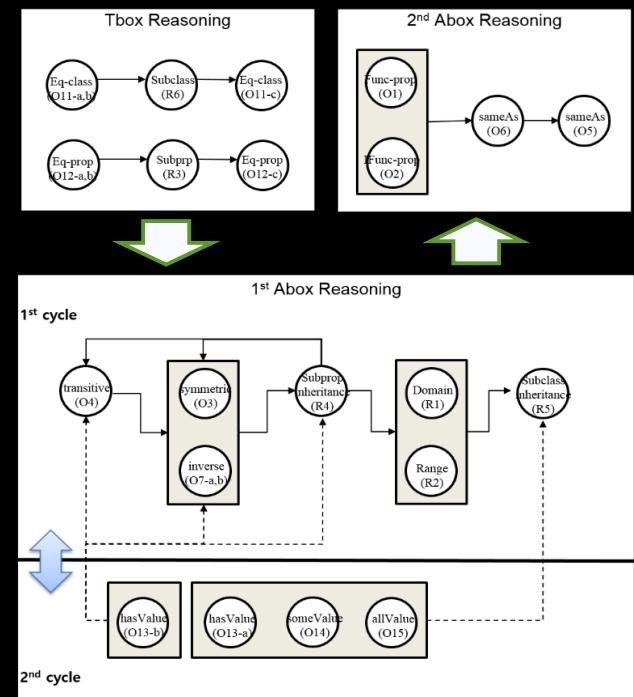


Spark based Parallel Reasoner

- World fastest OWL Reasoner working on Spark architecture – implementing DL rules on key-value data abstraction (Pair RDD)
- Optimization of reasoning sequence and recursive algorithm for improving the performance of OWL horst reasoning working on smaller memory



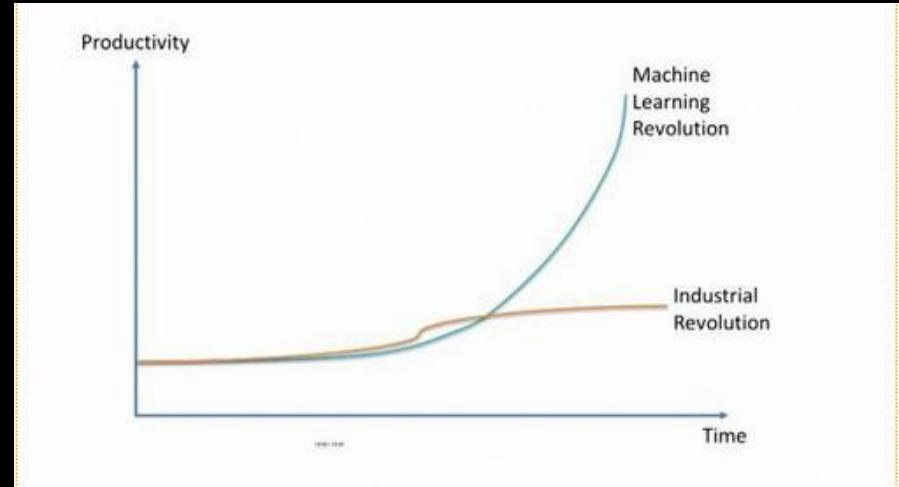
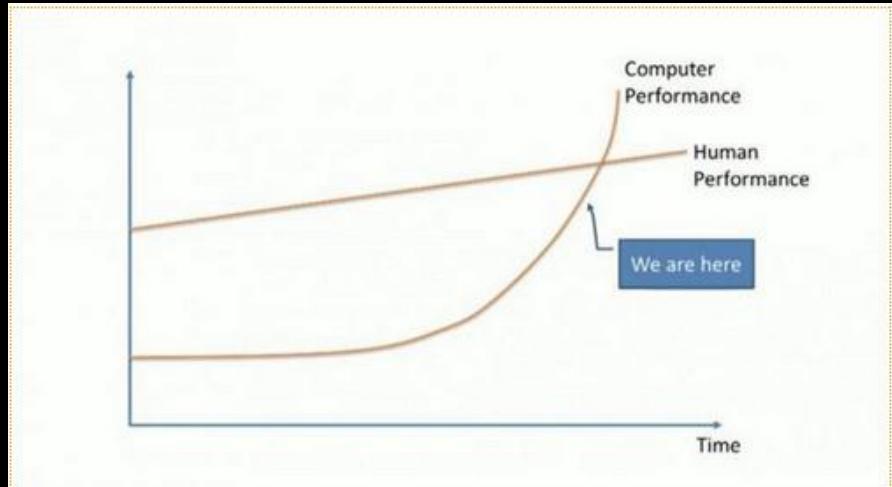
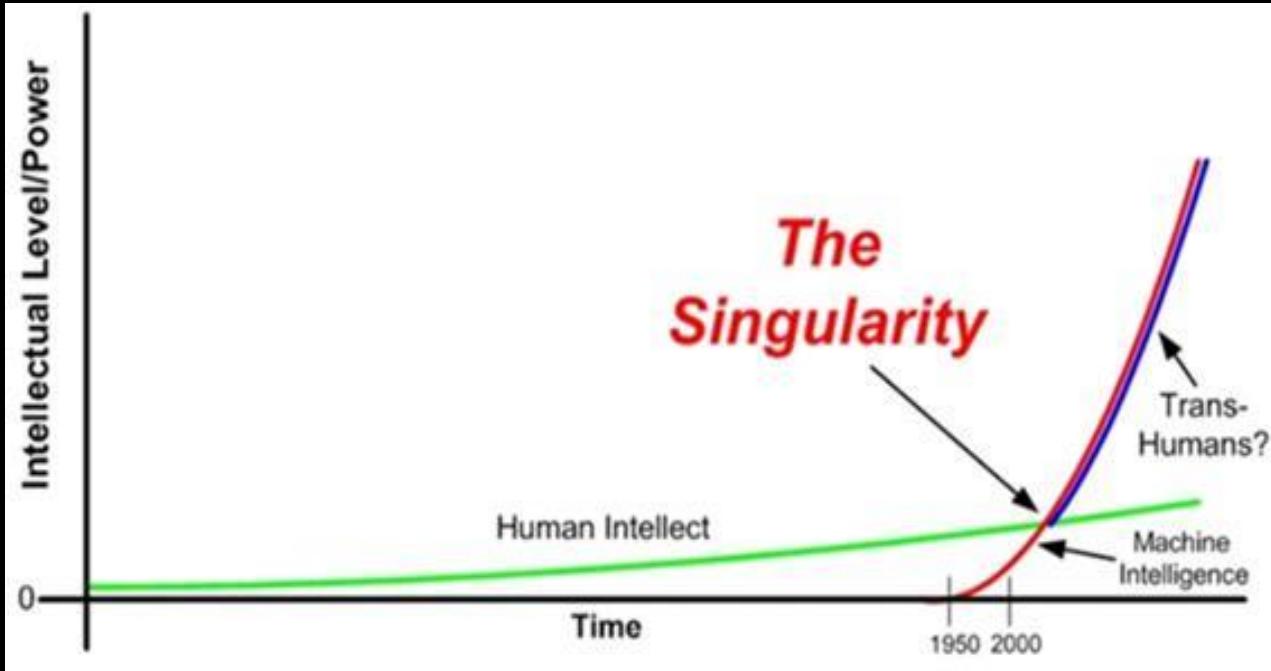
Configuration of Spark Reasoner



Optimization of reasoning sequence

Act Four

The Future of AI?



Intuition and Insight

Cost and Productivity

Enemy or Friend?

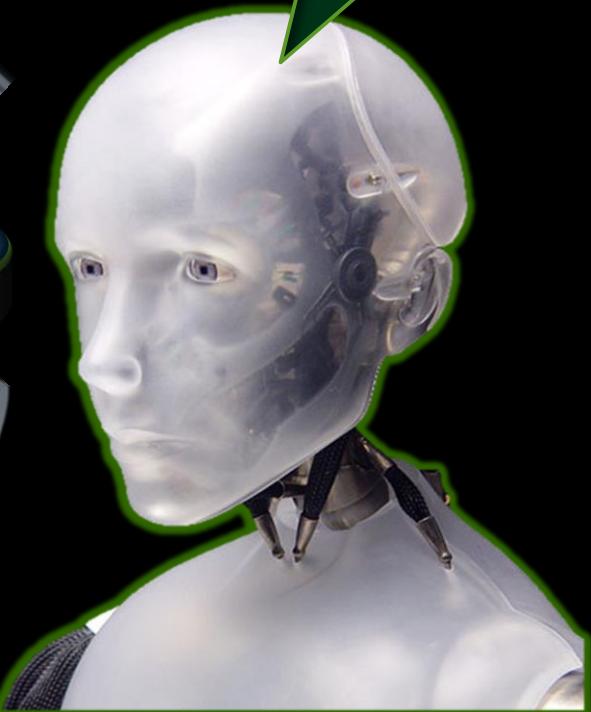
- Logics
- Rational

- Emotion
- Creativity

- Big data proc.
- Routine tasks



How they could
Collaborate?



Conclusion

“Computers are incredibly fast, accurate, and stupid.
Human beings are incredibly slow, inaccurate, and brilliant.
Together they are powerful beyond imagination.”

- Albert Einstein -

- The era of human and machine collaboration.
- Healthy goose rather than big golden egg.