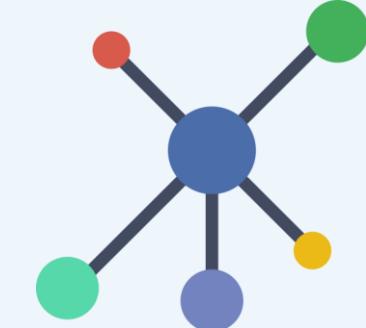


My Recent R&D Activities



At-a-glance Insight of Big Data



Akkharawoot Takhom, PhD

Research and Areas of Interest: Semantic Web and Knowledge Graph

Date and Time: Saturday, December 4, 2021; 10.00-11.00

Online Seminar via Microsoft Team:

* Ontology | * Collaboration | T Network Text Analysis

Bibliography

MFU

Since 2009: Internship student at NECTEC
•Human Language Technology Laboratory (**HLT**)



HLT, NECTEC

NECTEC

2009-11: Term of Reference at NECTEC
A Member of Text Processing Group, **HLT**
Language and Semantic Technology Laboratory (**LST**)



JAIST



SIIT

TAIST

SIIT-NECTEC Batch 4

2011-2013: M.Eng. of ICT for Embedded Systems
Thesis Topic: Ontology-enhanced Life Cycle Assessment (O-LCA): Toward formalizing the standard guidelines with a case study of application in oil refinery



LST, NECTEC

TAIST

JAIST

SIIT-NECTEC-JAIST
Batch 4

2014-2018 Ph.D. of Knowledge Science, JAIST

Dissertation Topic: Collaborative Development Approach for Multidisciplinary Ontology: A Scenario-based Knowledge Construction System in Life Cycle Assessment

2014 - 2019 Ph.D. of Doctor of Philosophy in Engineering and Technology, SIIT

Dissertation Topic: A Collaborative Ontology Development Framework for Enabling Multidisciplinary Knowledge Constructions



LCA, MTEC

NECTEC

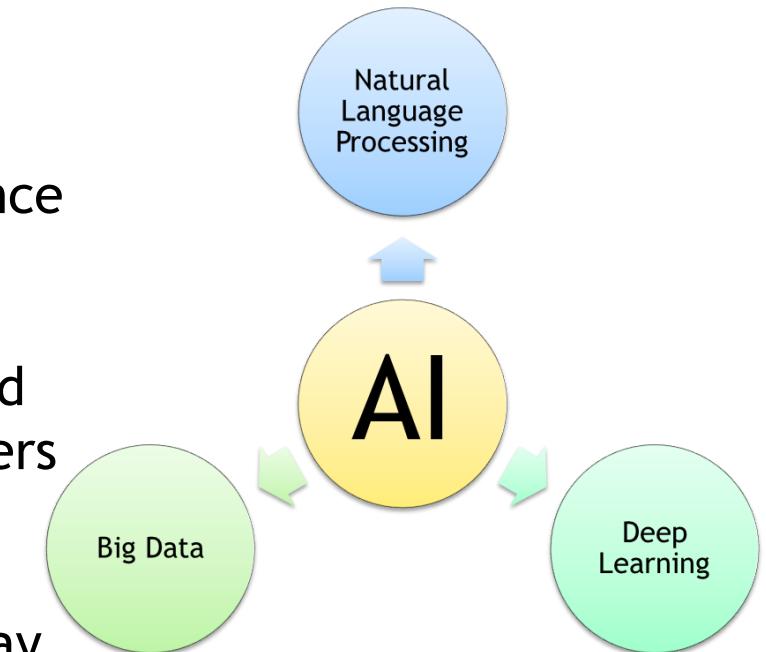
2018-present: Postdoctoral Fellowship Program

R&D Topic: Semantic Web/ Linked Data Migration Framework

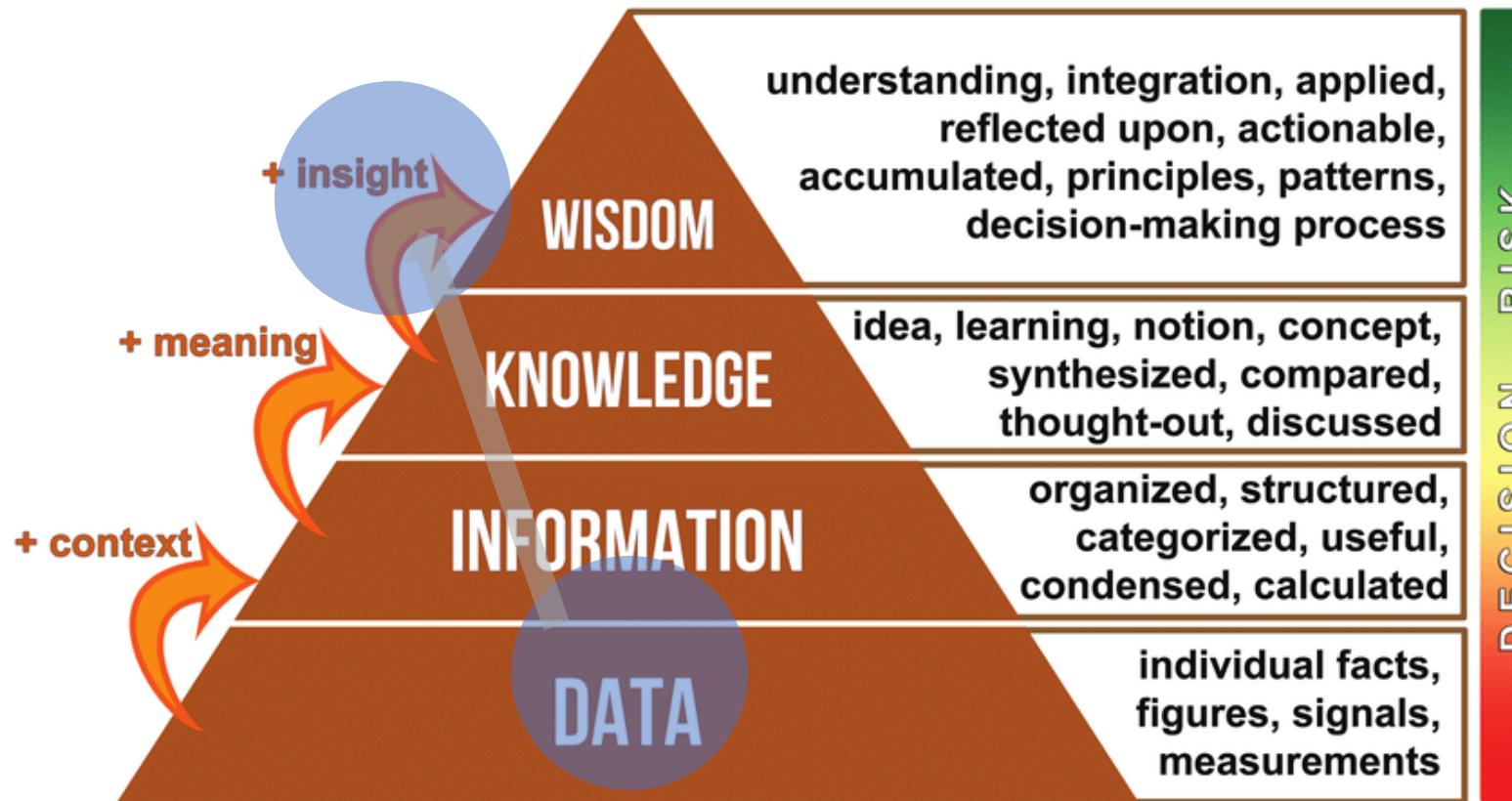


Artificial Intelligence (AI)

- The research and the foundations of **AI** date back to the 50's and 60's.
- The exponential pace at which AI has advanced over the past decade can be attributed to the proliferation of high-performance computing and the explosion of **Big Data**.
- High-Performance Computing is now ubiquitous, inter-connected networks and cloud computing allow **data scientists** and engineers to tackle complex problems and invent powerful solutions.
- **Big Data** is being generated all around us every second of the day and the rise of social media, smartphones, and the internet of things (IoT) have added sensors to everything from washing machines to oil tankers.



DIKW Pyramid



The **data-information-knowledge-wisdom (DIKW) hierarchy**

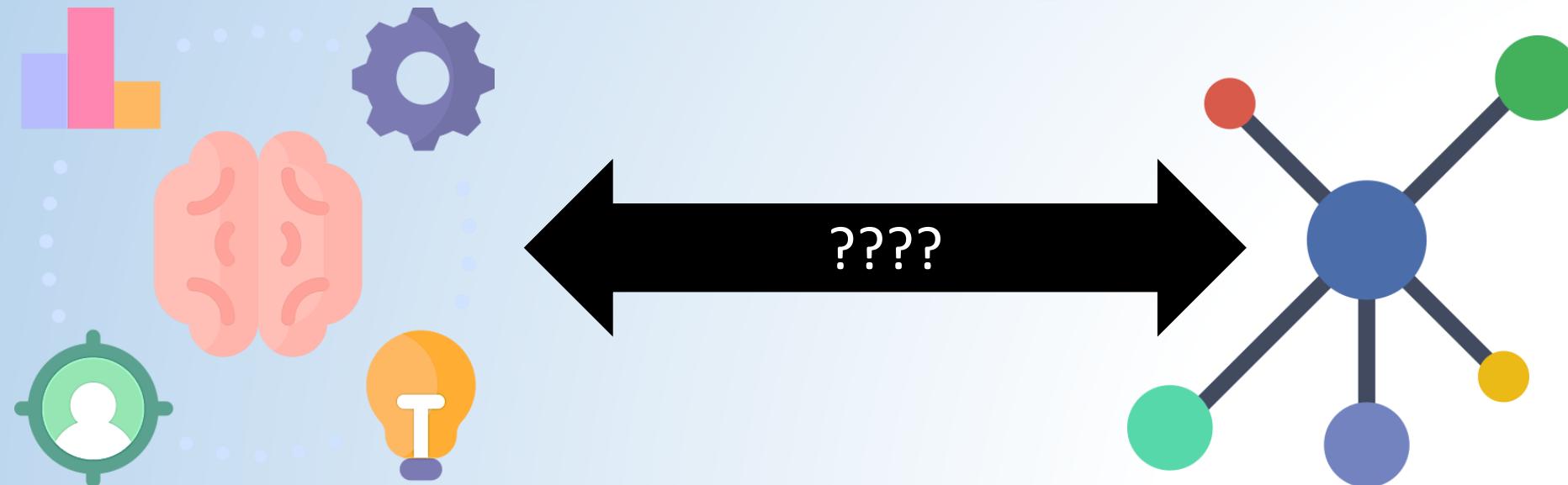
as a pyramid to manage knowledge.

Reproduced with permission from Luis Orlando Tedeschi (2019).

References:

- ASN-ASAS SYMPOSIUM: FUTURE OF DATA ANALYTICS IN NUTRITION: Mathematical modeling in ruminant nutrition: approaches and paradigms, extant models, and thoughts for upcoming predictive analytics.
Journal of Animal Science,
<https://academic.oup.com/jas/article/97/5/1921/5382308>

What is a Knowledge Graph?



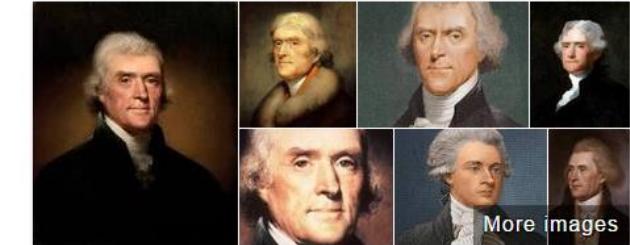
Knowledge Graph

Knowledge Graph is a knowledge base used by Google and its services to enhance its search engine's results with **information** gathered from a variety of sources.

- The **information** is presented to users in an **infobox** next to the search results.

Infoboxes were added to Google's search engine in May 2012, starting in the United States, with international expansion by the end of the year.

- Google has referred to these **infoboxes**, which appear to the right (top on mobile) of search results, as "knowledge panels"



Thomas Jefferson

3rd U.S. President

Thomas Jefferson was an American Founding Father, the principal author of the Declaration of Independence, and the third President of the United States. [Wikipedia](#)

Born: April 13, 1743, Shadwell, VA

Died: July 4, 1826, Charlottesville, VA

Presidential term: March 4, 1801 – March 4, 1809

Spouse: Martha Jefferson (m. 1772–1782)

Party: Democratic-Republican Party

Awards: AIA Gold Medal

[Get updates about Thomas Jefferson](#)

[Keep me updated](#)

People also search for

[View 15+ more](#)



John Adams



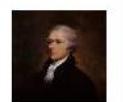
George Washington



Benjamin Franklin



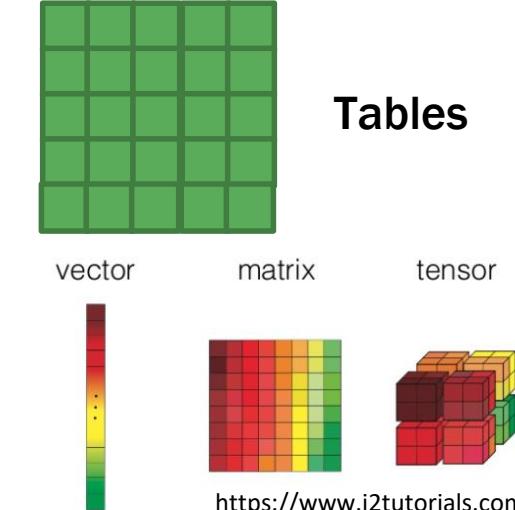
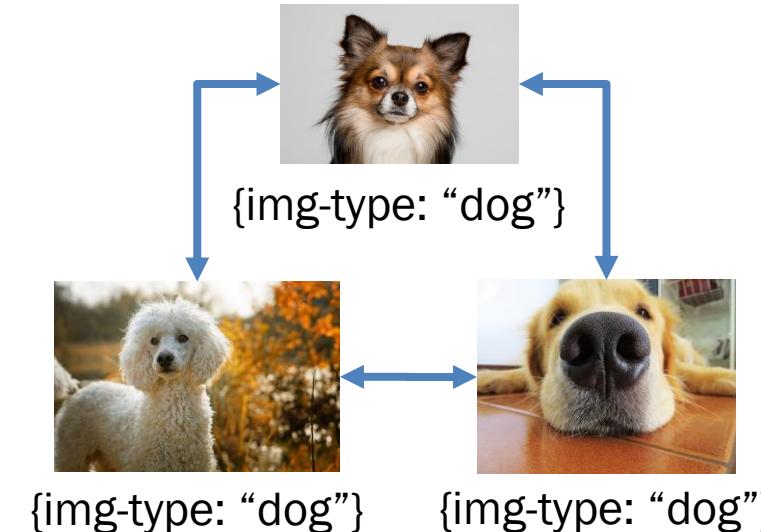
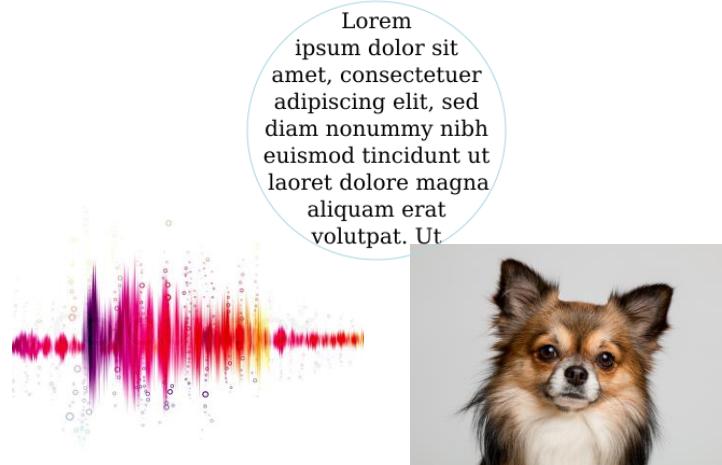
James Madison



Alexander Hamilton

Big Data describes the **vast amounts of data** - both structured and unstructured - that is now collected and stored each day.

Spectrum of Data (การจำแนกประเภทข้อมูล)



Unstructured Data

- Not having any pre-defined data model
- Hard to process due to ambiguity
- E.g. text, image, speech, and video

Semi-Structured Data

- Organized into an (over)-simplistic data model
- Associated with metadata for easy access
- E.g. semantic web

Structured Data

- Organized into a specific format (a.k.a. data model)
- Easy to access and process
- E.g. database, ontology

Example of Semi-Structured Data

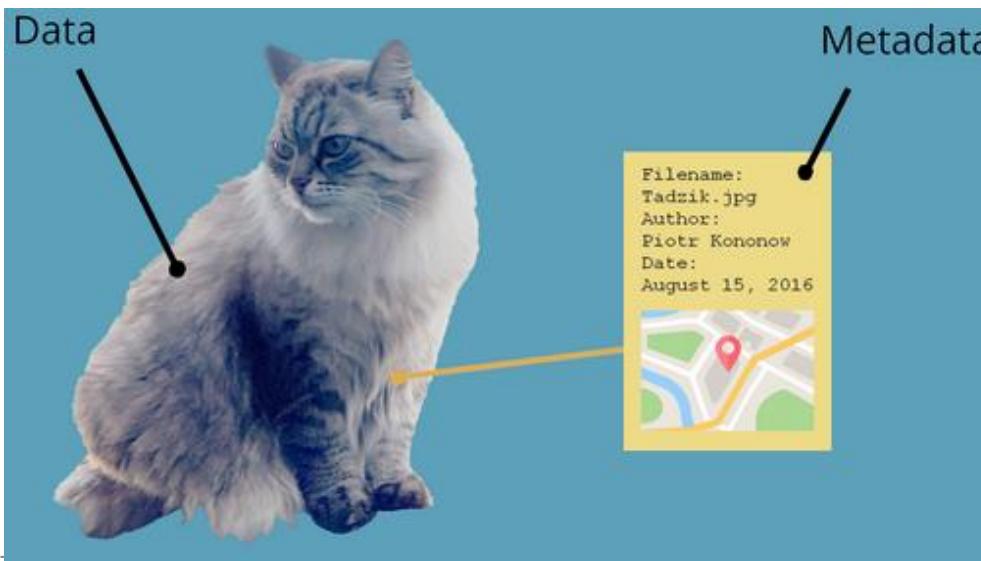
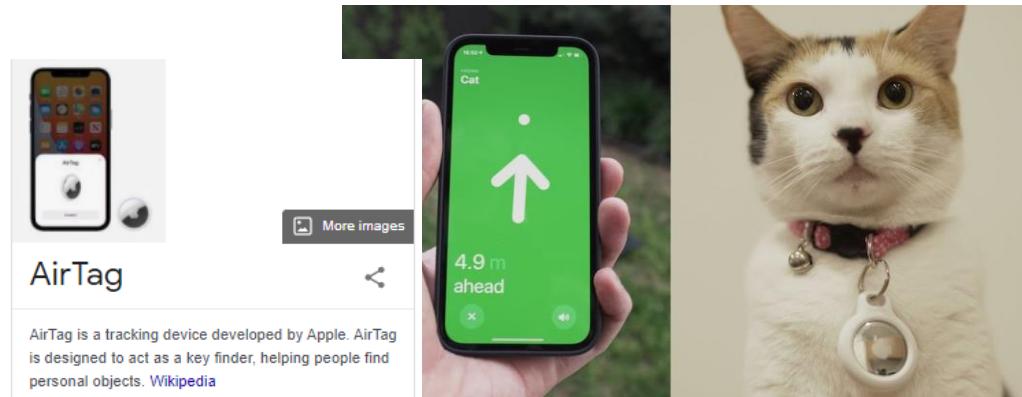
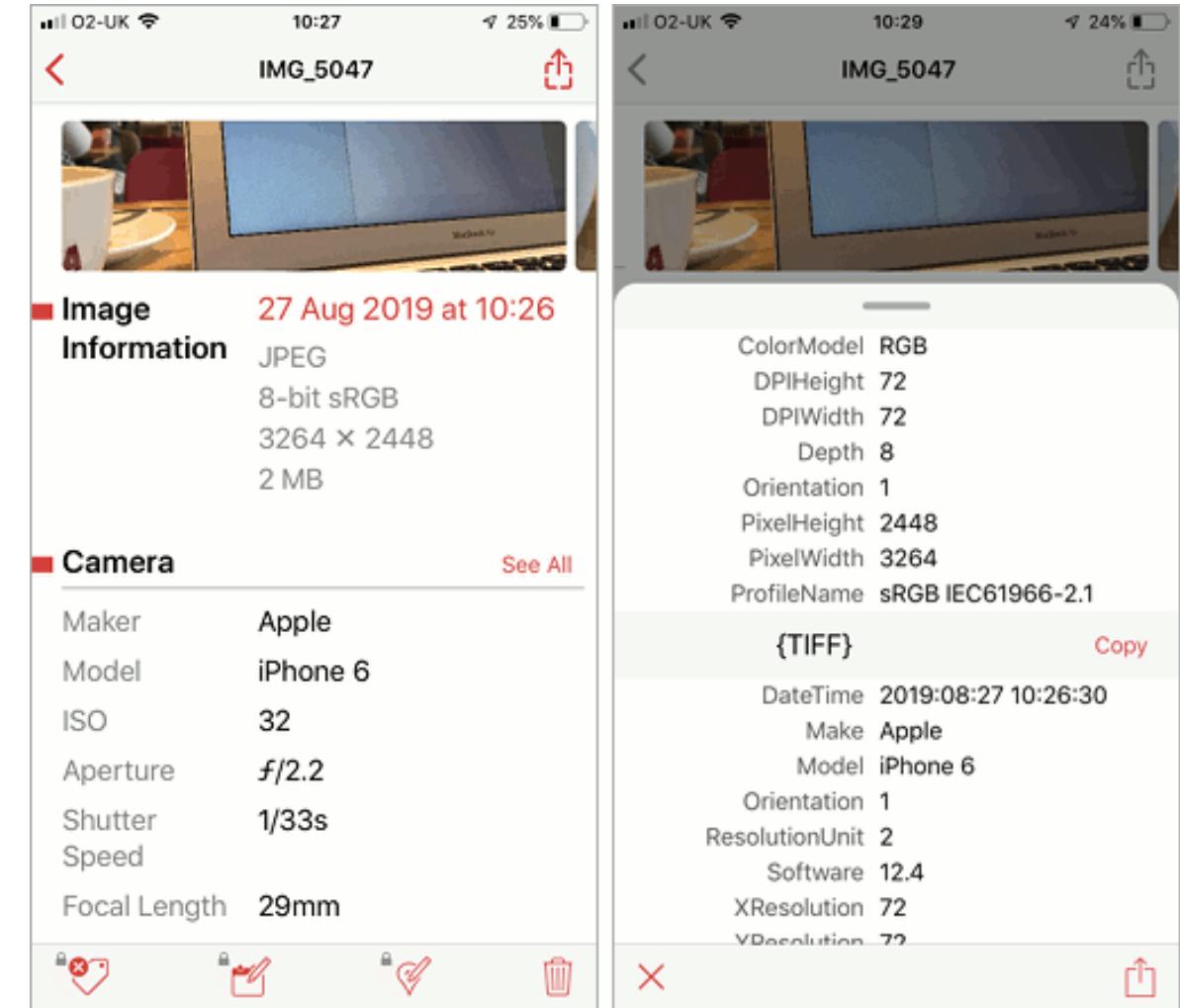



Image Information	
27 Aug 2019 at 10:26	ColorModel RGB
JPEG	DPIHeight 72
8-bit sRGB	DPIWidth 72
3264 × 2448	Depth 8
2 MB	Orientation 1

Camera	
See All	{TIFF} Copy
Maker	Date 2019:08:27 10:26:30
Model	Make Apple
ISO	Model iPhone 6
Aperture	Orientation 1
Shutter Speed	ResolutionUnit 2
Focal Length	Software 12.4

Example: Spectrum of Data & Knowledge Graph

Leonardo da Vinci^[5] (15 April 1452 – 2 May 1519) was an Italian polymath of the High Renaissance who was active as a painter, draughtsman, engineer, scientist, theorist, sculptor and architect.^[3] While his fame initially rested on his achievements as a painter, he also became known for his notebooks, in which he made drawings and notes on a variety of subjects, including anatomy, astronomy, botany, cartography, painting, and paleontology. Leonardo's genius epitomized the Renaissance humanist ideal,^[4] and his collective works compose a contribution to later generations of artists matched only by that of his younger contemporary, Michelangelo.^{[3][4]}

Born out of wedlock to a successful notary and a lower-class woman in, or near, Vinci, he was educated in Florence by the Italian painter and sculptor Andrea del Verrocchio. He began his career in the city, but then spent much time in the service of Ludovico Sforza in Milan. Later, he worked in Florence and Milan again, as well as briefly in Rome, all while attracting a large following of imitators and students. Upon the invitation of Francis I, he spent his last three years in France, where he died in 1519. Since his death, there has not been a time where his achievements, diverse interests, personal life, and empirical thinking have failed to incite interest and admiration,^{[3][4]} making him a frequent namesake and subject in culture.

Leonardo is among the greatest painters in the history of art and is often credited as the founder of the High Renaissance.^[3] Despite having many lost works and less than 25 attributed major works—including numerous unfinished works—he created some of the most influential paintings in Western art.^[3] His magnum opus, *The Mona Lisa*, is his best known work and often regarded as the world's most famous painting. *The Last Supper* is the most reproduced religious painting of all time and his *Vitruvian Man* drawing is also regarded as a cultural icon. In 2017, *Salvator Mundi*, attributed in whole or part to Leonardo,^[5] was sold at auction for US\$450.3 million, setting a new record for the most expensive painting ever sold at public auction.

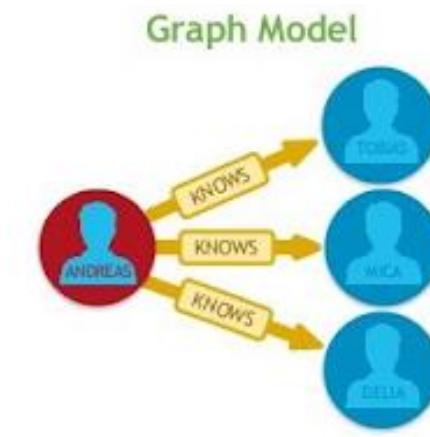
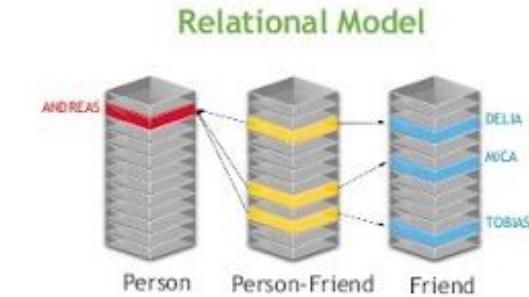
Revered for his technological ingenuity, he conceptualized flying machines, a type of armored fighting vehicle, concentrated solar power, an adding machine,^[6] and the double hull. Relatively few of his designs were constructed or even feasible during his lifetime, as the modern scientific approaches to metallurgy and engineering were only in their infancy during the Renaissance. Some of his smaller inventions, however, entered the world of manufacturing unheralded, such as an automated bobbin winder and a machine for testing the tensile strength of wire. He made substantial discoveries in anatomy, civil engineering, hydrodynamics, geology, optics, and tribology, but he did not publish his findings and they had little to no direct influence on subsequent science.^[7]

Unstructured Data



```
{name: "Leonardo da Vinci"}  
{notable_works: "Mona Lisa"}
```

Semi-Structured Data



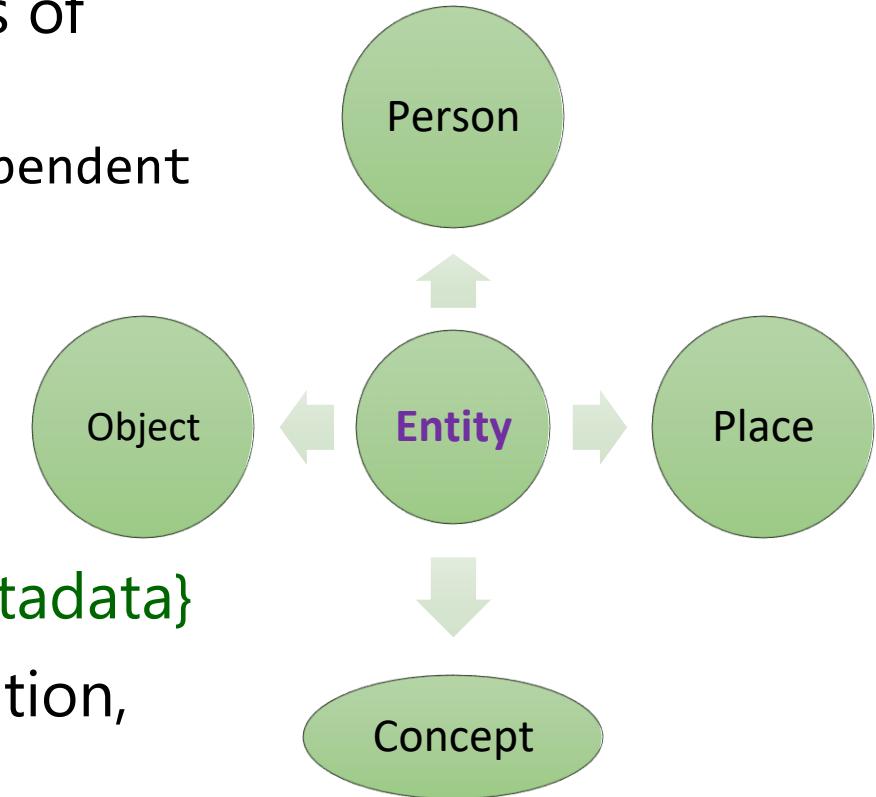
Structured Data

Knowledge Graph

- Represent a collection of interlinked descriptions of **entities**.

- Entity** (Noun) a thing with distinct and independent existence.
- Entity** is a class of persons/places/objects/events/concepts which we need to collect and store data

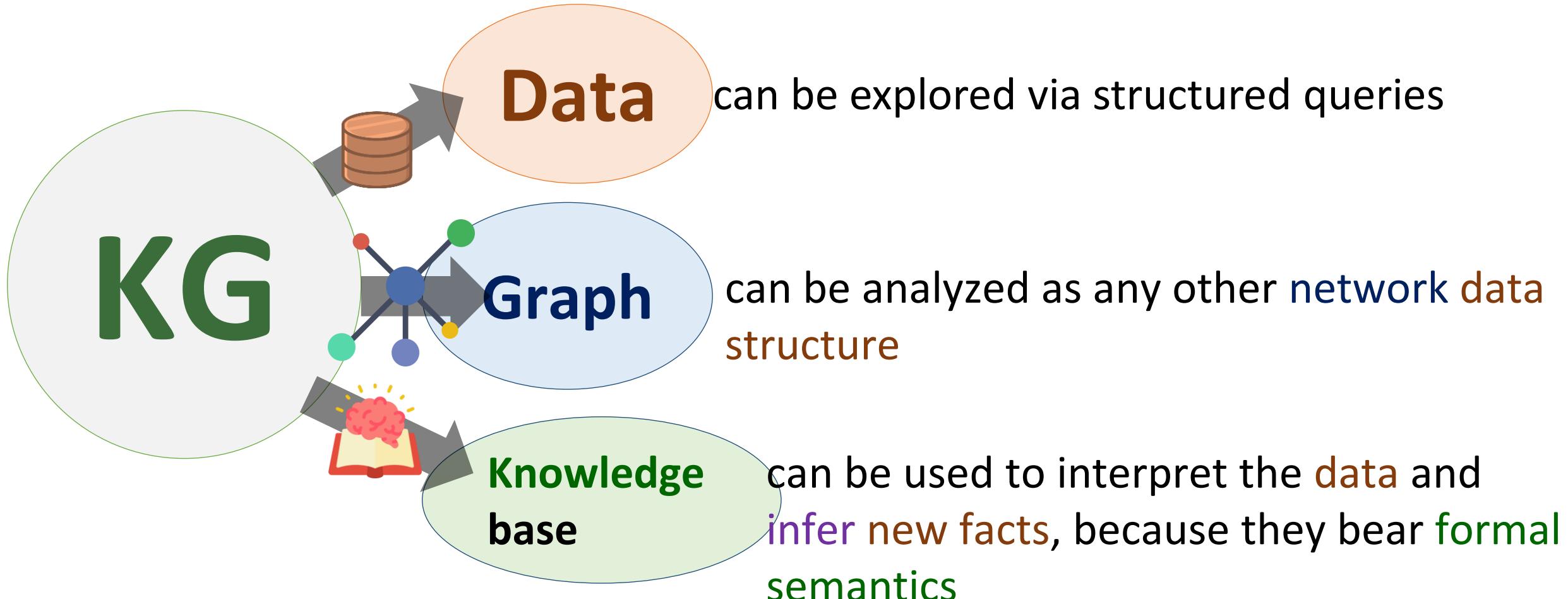
- Put data in **context** via linking and {semantic metadata}
- Provide a framework for data integration, unification, analytics and sharing.



Reference: What is a Knowledge Graph?

<https://www.ontotext.com/knowledgehub/fundamentals/what-is-a-knowledge-graph/>

Knowledge Graph: Key Characteristics



The Semantic Web

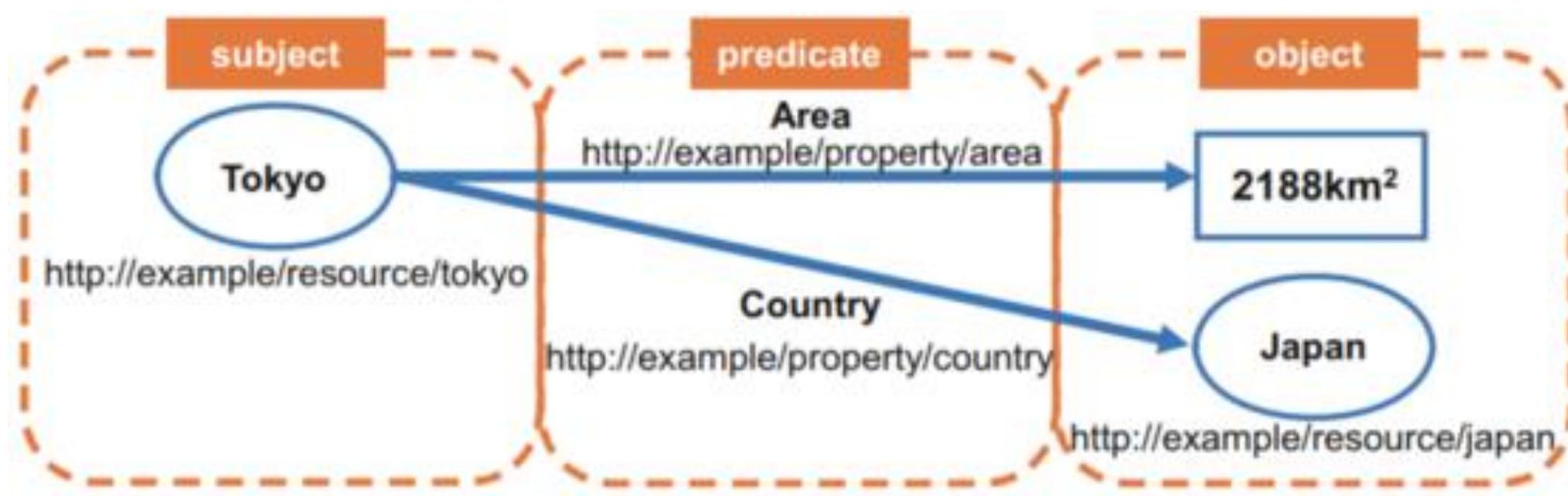
The Semantic Web is an W3C initiative to provide the data standards for data integration over the Web.

- Machine-readable and understandable data
- Structured and Linked data
- Uses global identifiers, i.e. **URI**, to refer to things

Resource Description Framework (RDF) is the core standard of the Semantic Web standards.

Triples

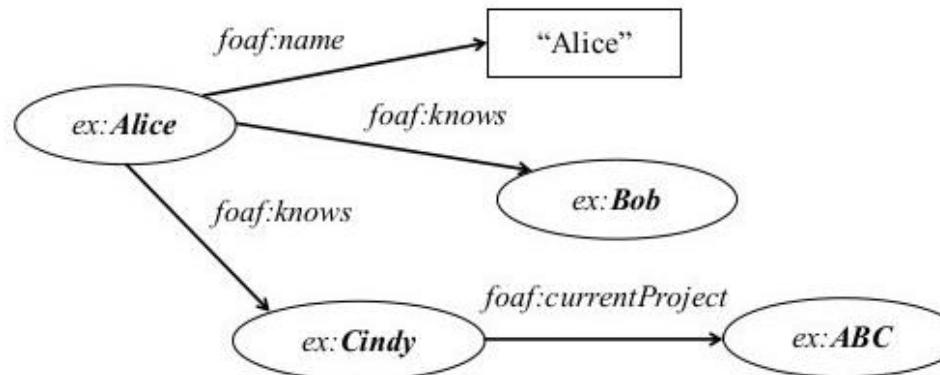
Triples connect entities to other entities to create a directed knowledge graph.



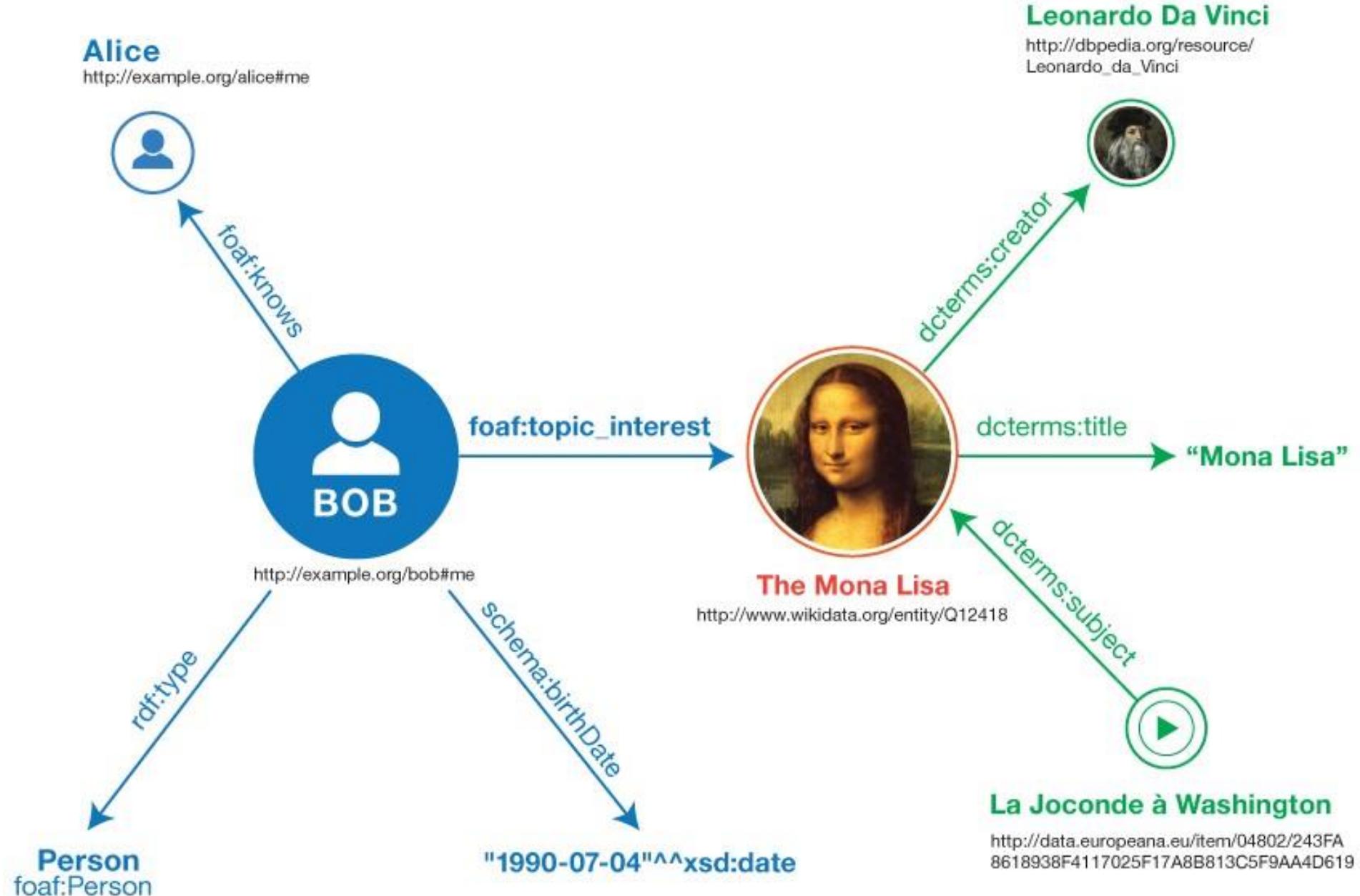
Friend of a Friend (FOAF)

FOAF is a machine-readable **ontology** describing **persons**, their activities and their relations to other **people** and objects.

- Anyone can use FOAF to describe themselves.
- Allows groups of people to describe social networks without the need for a centralized database.



ex:Alice	foaf:name	"Alice" .
ex:Alice	foaf:knows	ex:Bob .
ex:Alice	foaf:knows	ex:Cindy .
ex:Cindy	foaf:currentProject	ex:ABC .



Publications



Scopus : <https://www.scopus.com/authid/detail.uri?authorId=35231260200>

❖ [2018 Conference Paper] *A study of lexical ambiguity in large forum discussions for multidisciplinary knowledge engineering,*

- Takhom, A., Boonkwan, P., Ulrich Hoppe, H., ...Usanavasin, S., Supnithi, T.; The13th International Conference on Knowledge, Information and Creativity Support Systems, Proceedings, 2018.

❖ [2019 Conference Paper] *A Supportive Environment for Knowledge Construction based on Semantic Web Technology: A Case Study in a Cultural Domain,*

- Takhom, A., Leenoi, D., Soomjinda, P., Boonkwan, P., Supnithi, T., The14th International Joint Symposium on Artificial Intelligence and Natural Language Processing, iSAI-NLP 2019.

❖ [2020 Article] *Discovering cross-disciplinary concepts in multidisciplinary context through collaborative framework*

- Takhom, A., Usanavasin, S., Supnithi, T., Hoppe, H.U., Boonkwan, P. ; The International Journal of Knowledge and Systems Sciences, 2020, 11(2), pp. 1–19

❖ [2020 Article] *A collaborative framework supporting ontology development based on agile and scrum model*

- Takhom, A., Usanavasin, S., Supnithi, T., Boonkwan, P.; IEICE Transactions on Information and Systems, 2020, E103D(12), pp. 2568–2577

❖ [2020 Conference Paper] *An Approach of Network Analysis Enhancing Knowledge Extraction in Thai Newspapers Contexts*

- A. Takhom, D. Leenoi, C. Sophaken, P. Boonkwan, and T. Supnithi; The 2nd joint Workshop on NLP/AI R&D Workshop

❖ [2021 Article] *An Approach of Network Analysis Enhancing Knowledge Extraction in Thai Newspapers Contexts*

- A. Takhom, D. Leenoi, C. Sophaken, P. Boonkwan, and T. Supnithi; Journal of Intelligent Informatics and Smart Technology (Volume 6, October 2021)

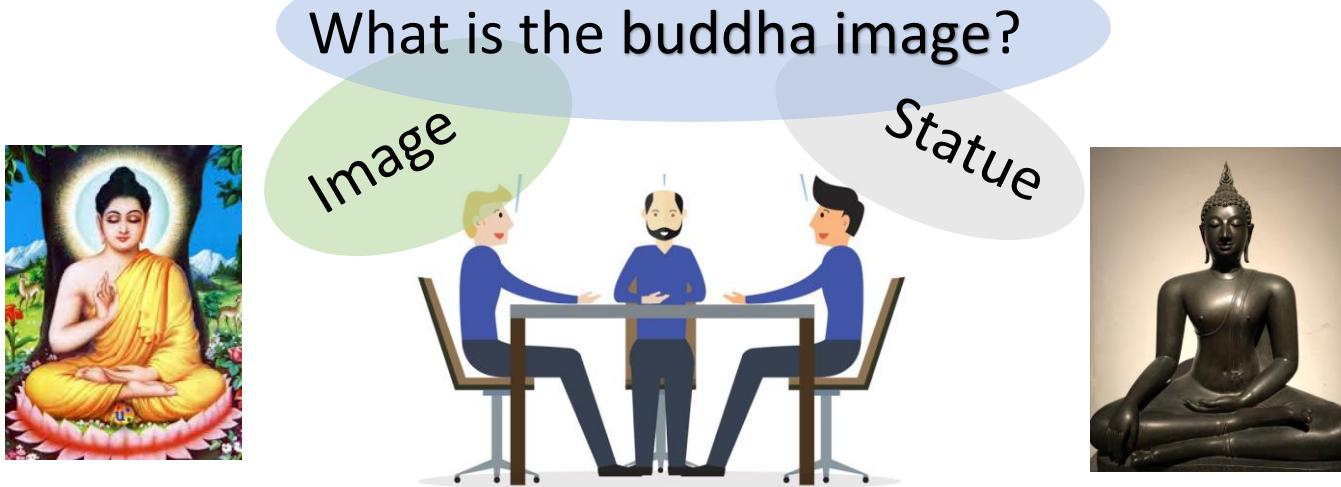
❖ [2021 Conference Paper] *Knowledge Graph Enhanced Community Consensus: a Scenario-based Knowledge Construction on Buddha Images*

- A. Takhom, T. Utasri, D. Leenoi, P. Soomjinda, P. Boonkwan and T. Supnithi.; The 10th International Joint Conference on Knowledge Graphs (IJCKG 2021)

[2019 Conference Paper]

A Supportive Environment for Knowledge Construction based on Semantic Web Technology: A Case Study in a Cultural Domain,

Takhom, A., Leenoi, D., Soomjinda, P., Boonkwan, P., Supnithi, T., The 14th International Joint Symposium on Artificial Intelligence and Natural Language Processing, iSAI-NLP 2019.



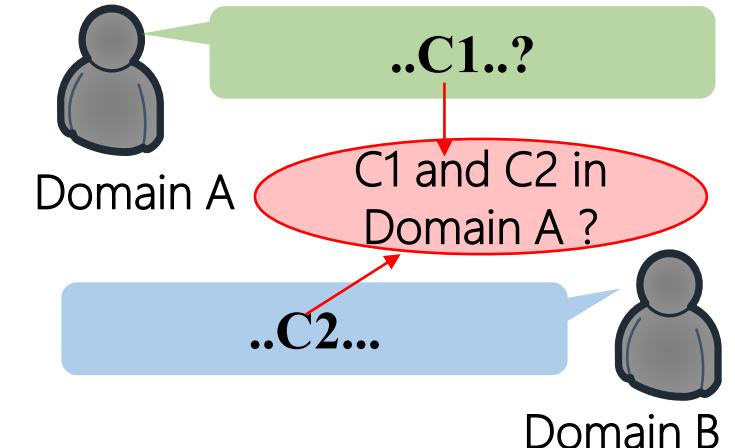
Semantic Web Technology [G. Stephan+, 2007]

- A crucial role in knowledge sharing within the ontological engineering and a specific-domain community.

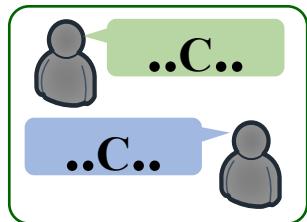
Knowledge Construction [A. Takhom+, 2018]

- To motivate many research works and modeling efforts of stakeholders, e.g., experts in particular knowledge

Lexical Ambiguity Problem



A Communicative Function



Supporting
Knowledge Construction

[2019 Conference Paper]

A Supportive Environment for Knowledge Construction based on Semantic Web Technology: A Case Study in a Cultural Domain,
Takhom, A., Leenoi, D., Soomjinda, P., Boonkwan, P., Supnithi, T., The14th International Joint Symposium on Artificial Intelligence and Natural Language Processing, iSAI-NLP 2019.

- We organized a collaborative workshop at Chiang Mai International Convention and Exhibition Center on February 2nd, 2019.
 - To encourage experts in specific domains to understand and work with the ontology-based Semantic Web technology.
- **Knowledge Representation & Sharing for a Cultural Domain**



Thai hill tribes



Thai ancient wagon



Thai Worship



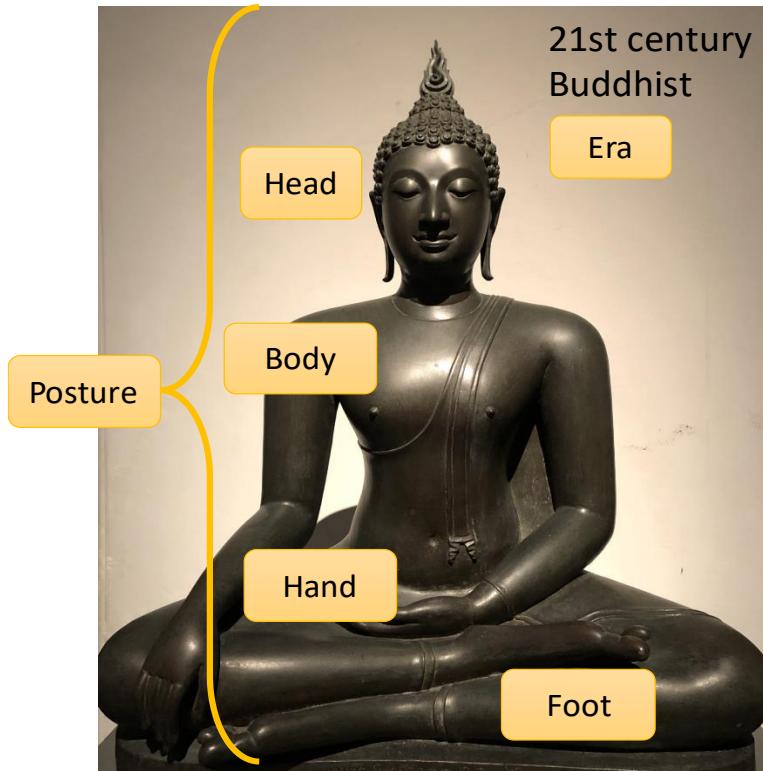
Buddha image

[2019 Conference Paper]

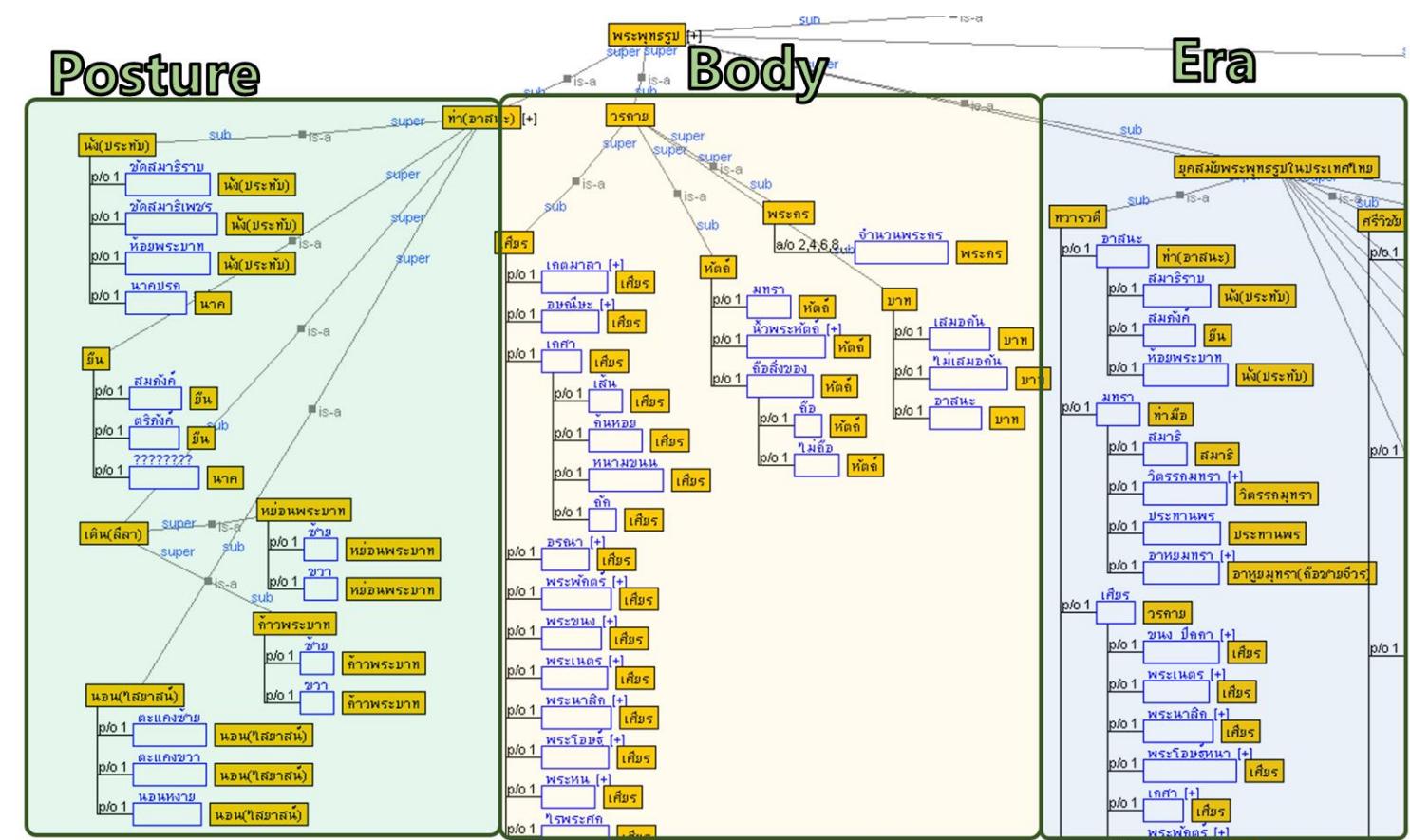
A Supportive Environment for Knowledge Construction based on Semantic Web Technology: A Case Study in a Cultural Domain,

Takhom, A., Leenoi, D., Soomjinda, P., Boonkwan, P., Supnithi, T., The14th International Joint Symposium on Artificial Intelligence and Natural Language Processing, iSAI-NLP 2019.

• Buddha Image's Ontology Version 1



A sample of the bronze Buddha statue in Sukhothai style, invented in 21st century.



[2020 Conference Paper] An Approach of Network Analysis Enhancing Knowledge Extraction in Thai Newspapers Contexts

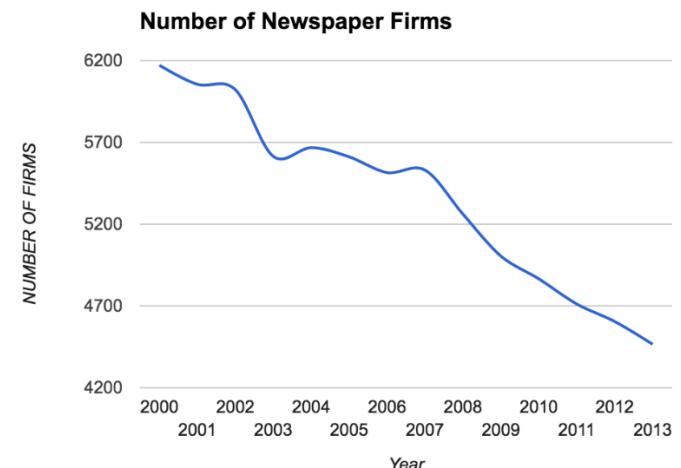
A. Takhom, D. Leenoi, C. Sophaken, P. Boonkwan, and T. Supnithi; The 2nd joint Workshop on NLP/AI R&D Workshop

Multidisciplinary Knowledge [Bernard & Anita, 2006]

- Several domains with possibly **incompatible** interpretations by particular domain experts.
- **Different** understandings when using the identical terms may:
 → **Obstruct** the sharing knowledge collaborating among dissimilar stakeholders
 → Lead to **miscommunication**,

Disruption in Information Technology

- Caused **newspapers** experiencing a significant decline.
- Relating to the advent of internet, not only the printed news having been challenged.
- BUT the **readers' knowledge acquisition has been changed.**

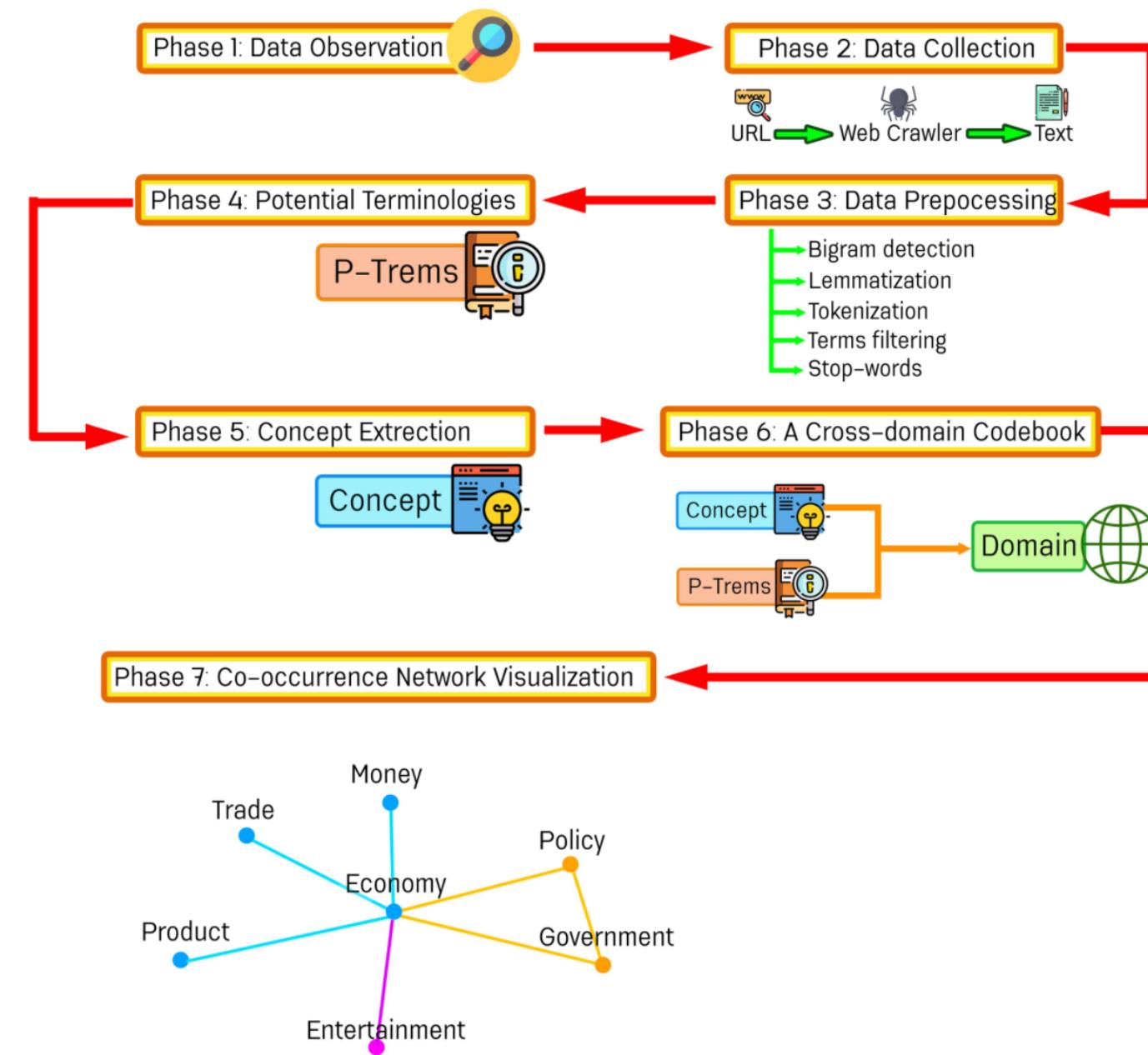


IT Disruption



The Network-Text Analysis workflow

21



7 Typical Phases of the NTA Workflow

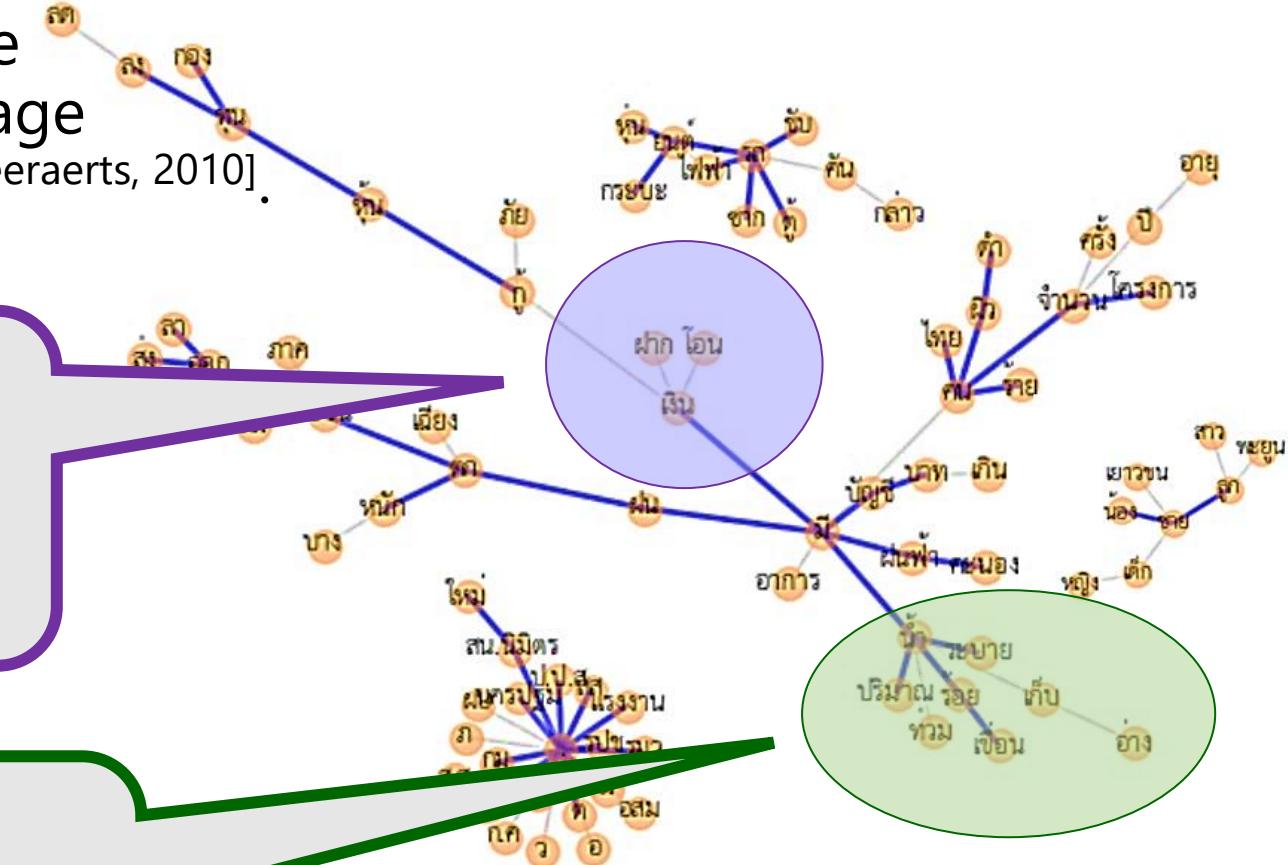
- Phase 1: Data observation
- Phase 2: Data collection
- Phase 3: Data preprocessing
- Phase 4: Potential Terminologies
- Phase 5: Concept extraction
- Phase 6: Building of cross-domain codebooks
- Phase 7: Visualization of co-occurrence networks

Result and Discussion

The figure is a group of words which evoke readers' mind to the Frame one of the desiderata for natural language processing: **FrameNet** [Fillmore, 1985, 2008; Geeraerts, 2010].

'ฝาก' /fà:k/ 'deposit',
 'โอน' /?o:n/ 'transfer' and
 'เงิน' /t?u:n/ 'money' could evoke to the banking scenario in **finance domain**,

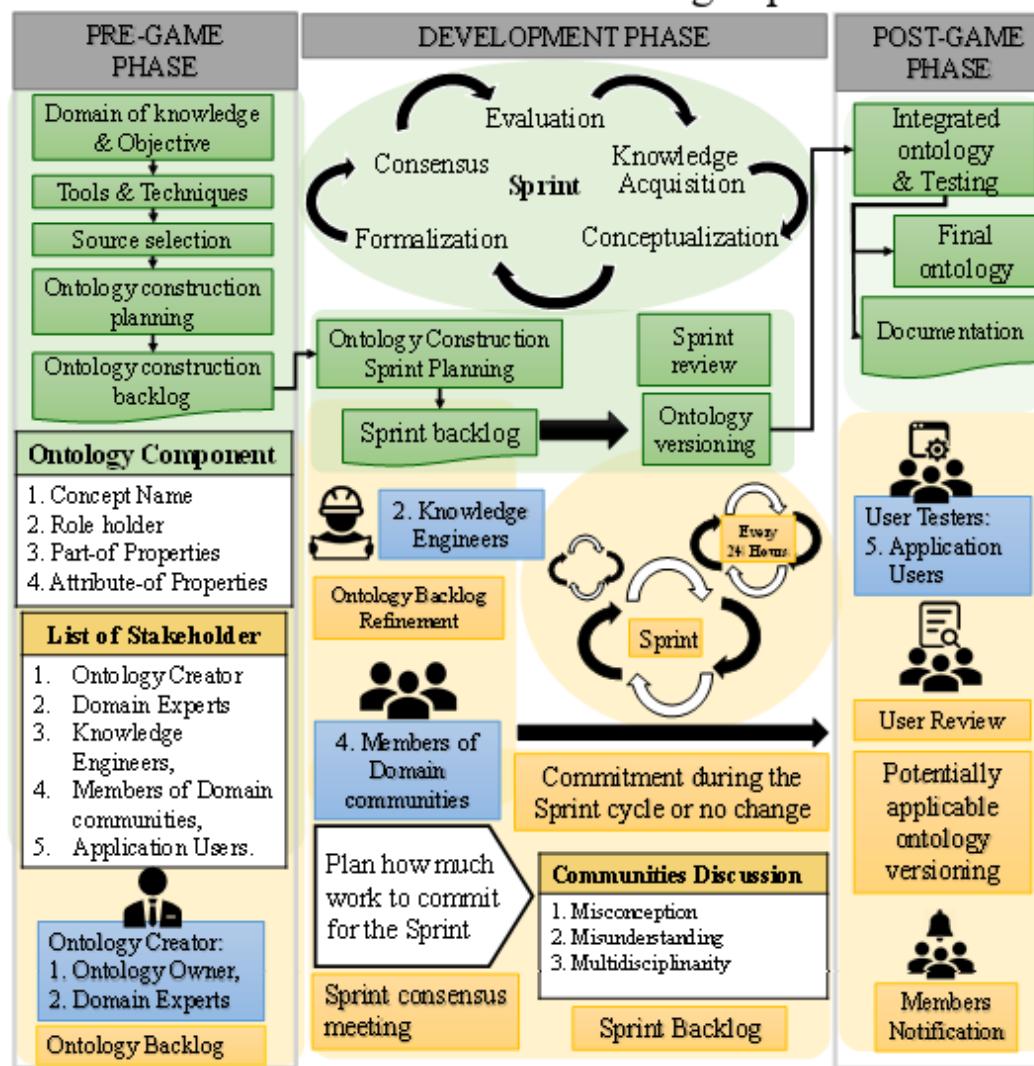
'อ่าง-เก็บ-น้ำ' /?à:tŋ kèp ná:m/ 'reservoir',
 'ระบายน้ำ' /ra? ba:i/ 'drain',
 'ท่วม' /tʰú:am/ 'flood' and 'เขื่อน' /kʰù:an/ 'dam'
 could lead to the **environment domain**



A generated co-occurrence network of multidisciplinarity.

[2020 Article] A collaborative framework supporting ontology development based on agile and scrum model

Takhom, A., Usanavasin, S., Supnithi, T., Boonkwan, P.; IEICE Transactions on Information and Systems this link is disabled, 2020, E103D(12), pp. 2568–2577





CD-OAM: Community-Driven Ontology-based Application Management Supporting Agile Ontology Development based on Scrum Model and Hozo Ontology Editor

Features: * Ontology | * Collaboration | * Network Text Analysis

Ontology

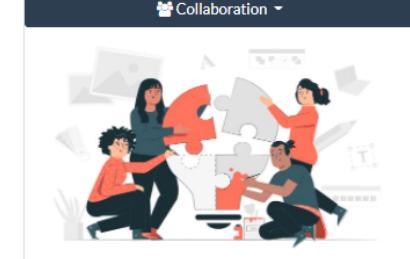


A Collaborative Framework Supporting Ontology Development based on Agile and Scrum Model

Journal: Akkharawoot et al. 2019

Ontology describes concepts and relations in a specific domain-knowledge that are important for knowledge representation and knowledge sharing. We incorporated ontology development process into Scrum process as used for process standard in software engineering. [Click for more details]

Collaboration



Collaborative Ontology Development Approach for Multidisciplinary Knowledge: A Scenario-based Knowledge Construction System in Life Cycle Assessment

Journal: Akkharawoot et al. 2018

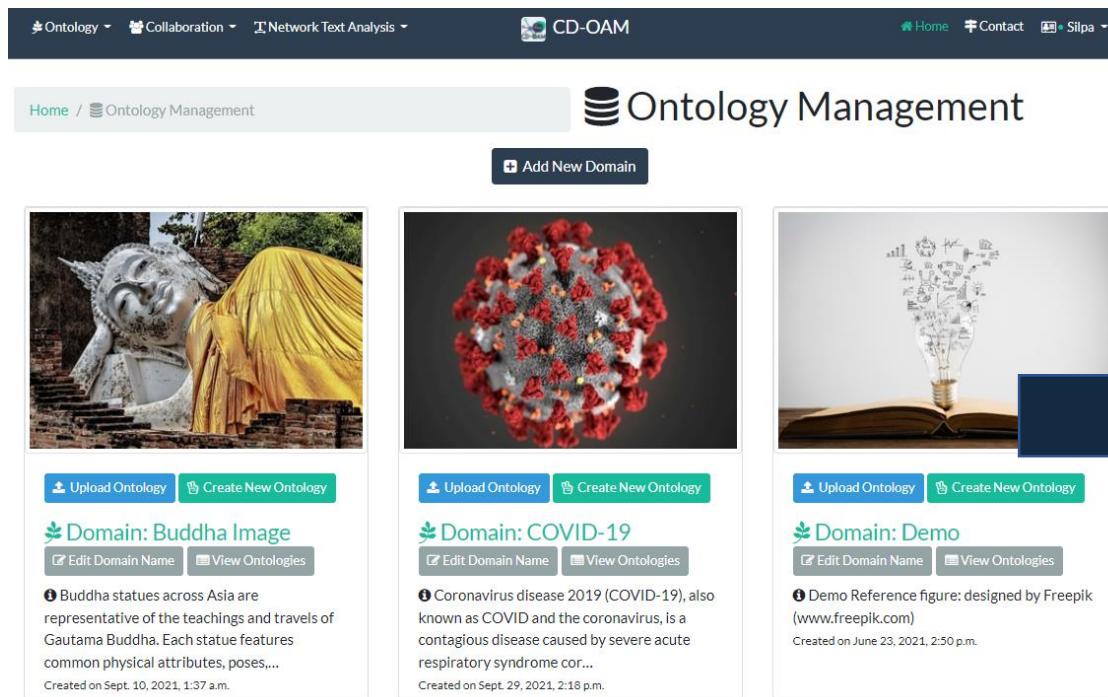
The community-driven ontology-based application management (CD-OAM) framework provides a collaborative environment with supporting features to enable collaborative knowledge creation. [Click for more details]

Community-Driven Ontology-based Application Management (CD-OAM)
<https://kgnlp.language-semantic.org/cdoam/>

[2021 Conference Paper]

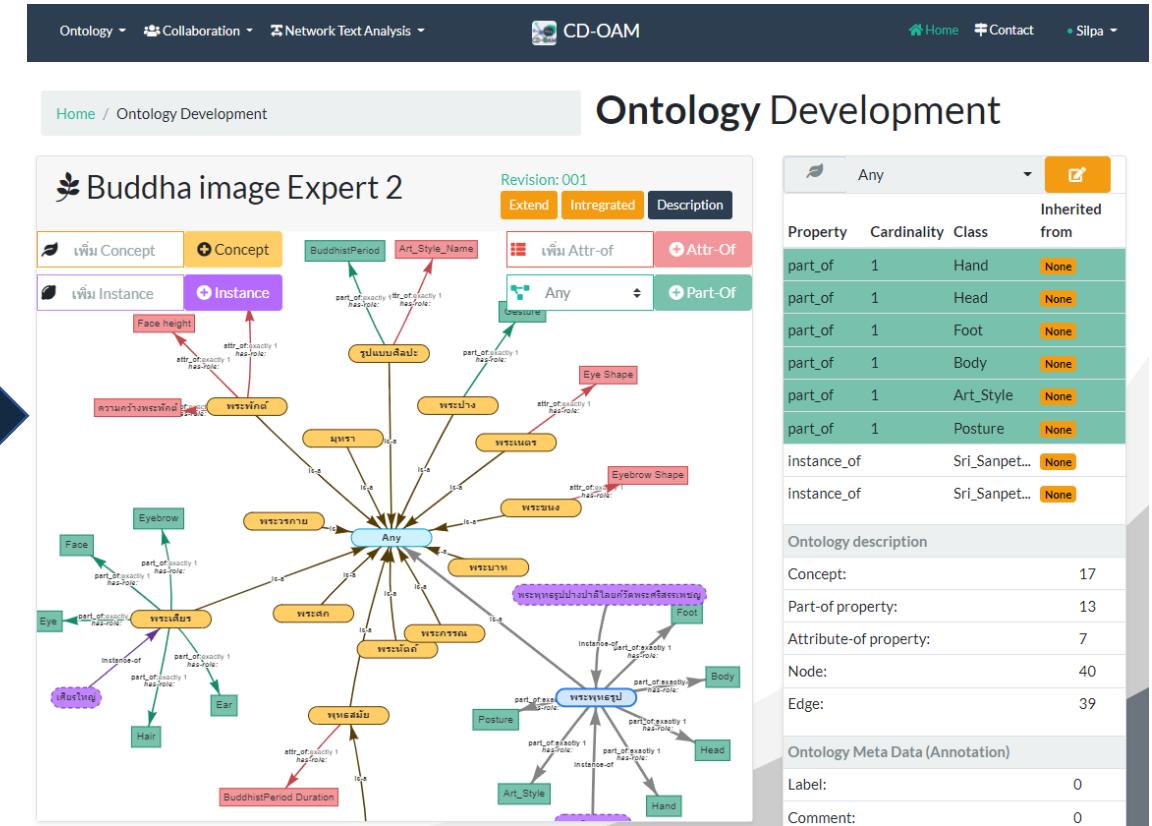
Knowledge Graph Enhanced Community Consensus: a Scenario-based Knowledge Construction on Buddha Images

A. Takhom, T. Utasri, D. Leenoi, P. Soomjinda, P. Boonkwan and T. Supnithi.; The 10th International Joint Conference on Knowledge Graphs (IJCKG 2021)



Ontology Management

- Buddha Image:** Domain: Buddha Image. Description: Buddha statues across Asia are representative of the teachings and travels of Gautama Buddha. Each statue features common physical attributes, poses... Created on Sept. 10, 2021, 1:37 a.m.
- COVID-19:** Domain: COVID-19. Description: Coronavirus disease 2019 (COVID-19), also known as COVID and the coronavirus, is a contagious disease caused by severe acute respiratory syndrome cor... Created on Sept. 29, 2021, 2:18 p.m.
- Demo:** Domain: Demo. Description: Demo Reference figure: designed by Freepik (www.freepik.com) Created on June 23, 2021, 2:50 p.m.



Future Work

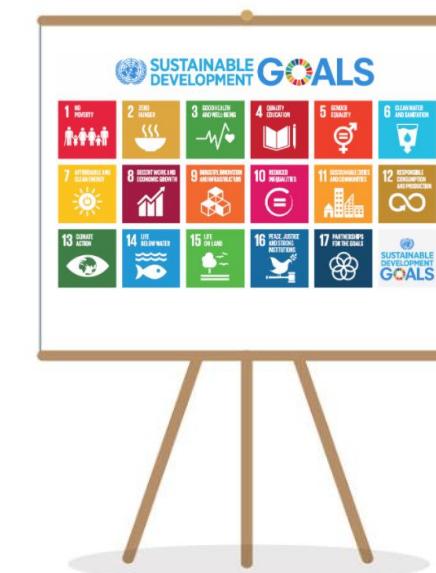
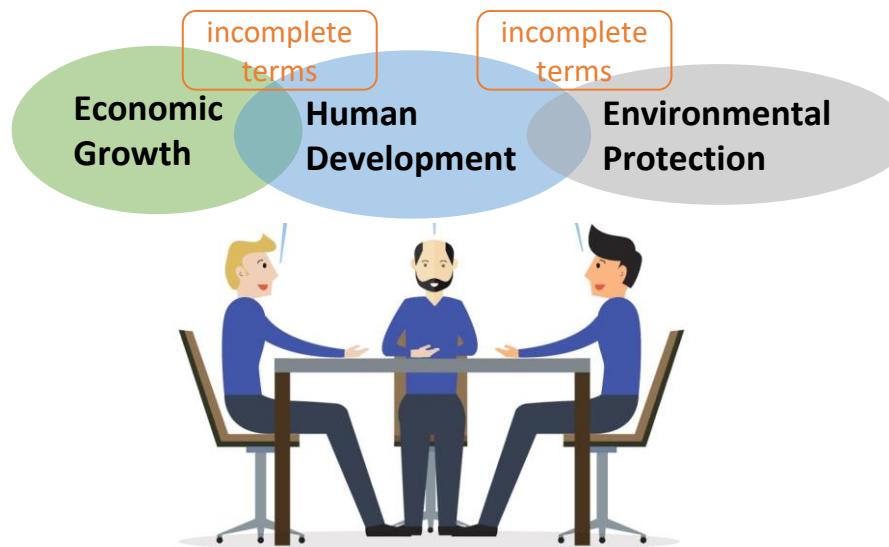
An understanding of characteristics of collaboration has authoritative in influencing stakeholder achieving practical research goals.

- **Blind Spots of Communication:**

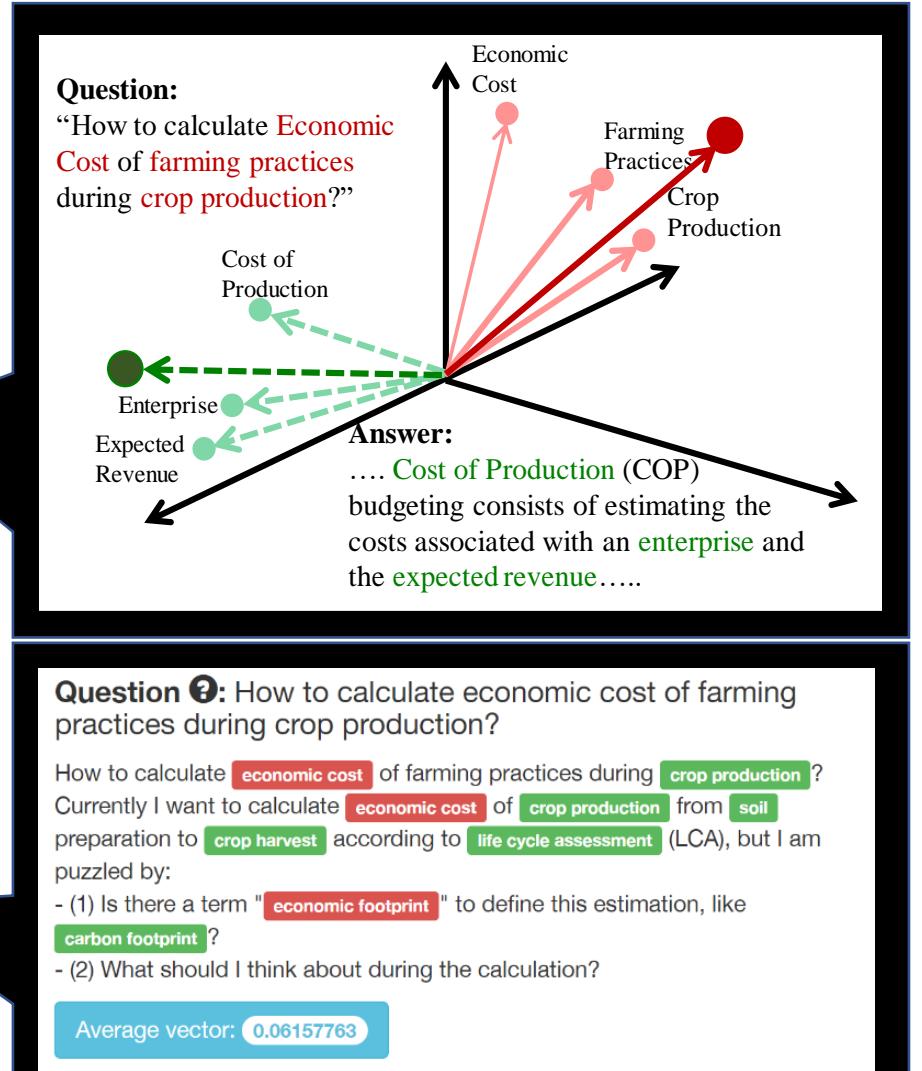
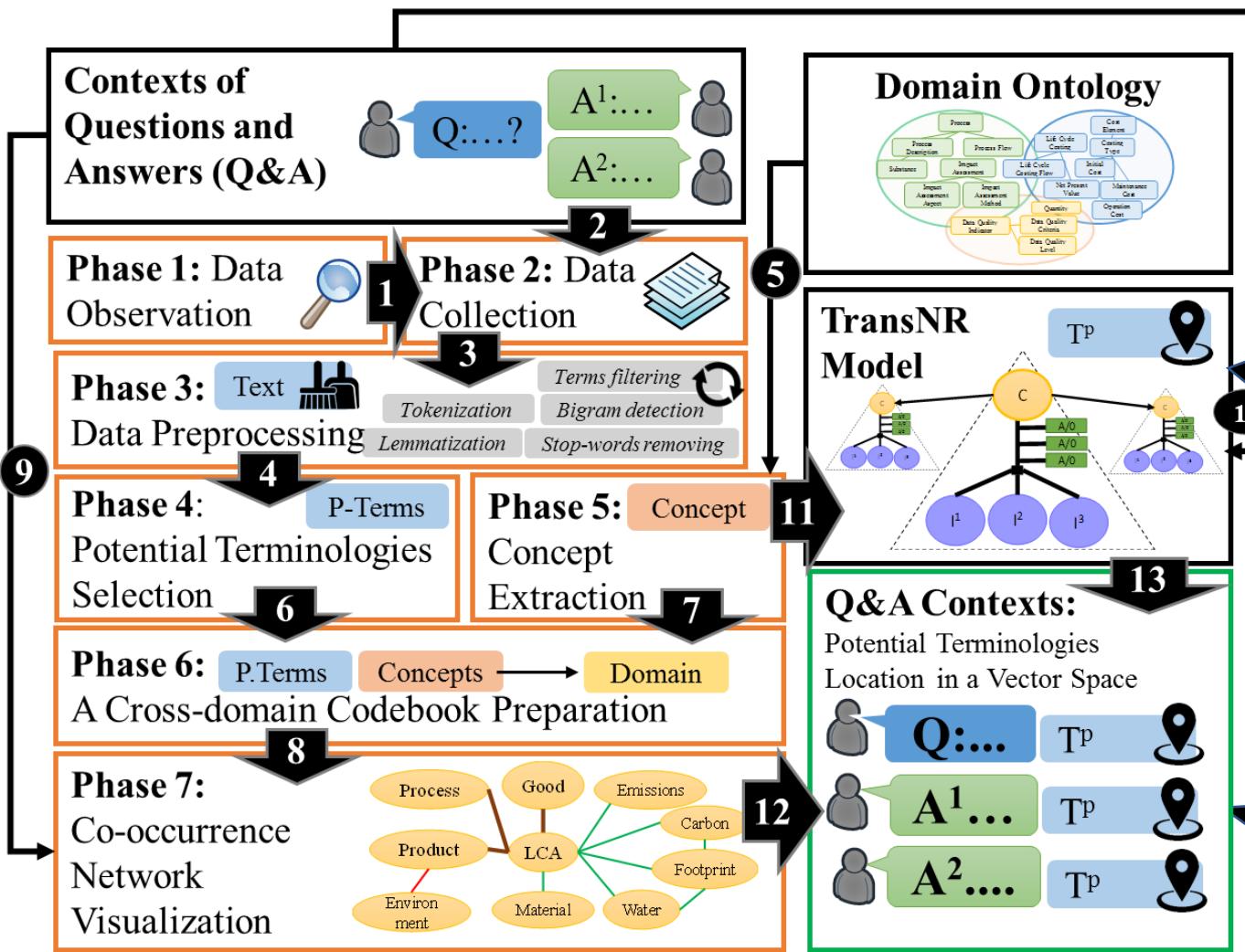
An interaction between stakeholders could be difficult when information is not correct by incomplete terms.

- **Multidisciplinary Knowledge:**

Multiple disciplines existing in sustainability science [2] involving



Future Work



ACKNOWLEDGMENT

This research is partially supported by

1. National Electronics and Computer Technology Center (NECTEC), Thailand.
2. Princess Chulabhorn Science High School (PCSHS), Thailand.
3. Faculty of Fine Arts, Chiang Mai University, Thailand

