Seminar: Knowledge Graphs

Alsayed Algergawy

Second week: 26.04.2023

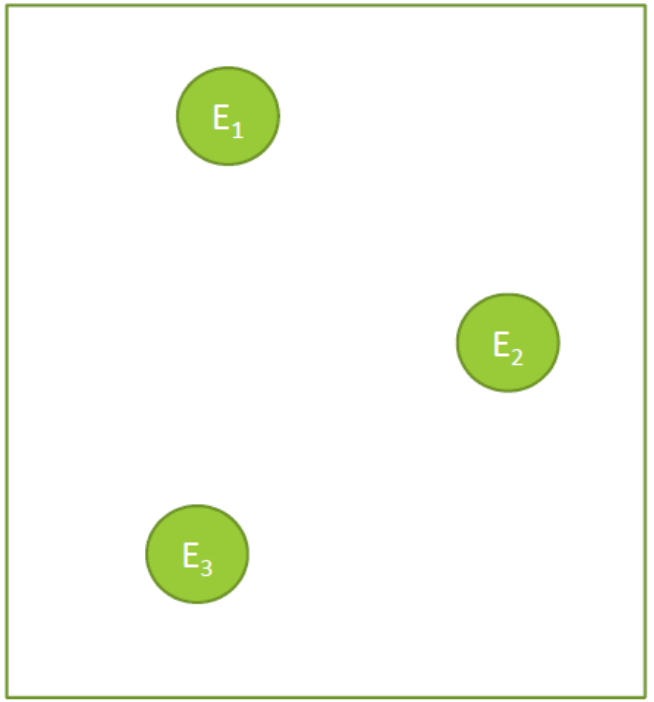
Module: 5369S

Agenda

- Recap
- Topics: An overview
- Topic allocation
- Seminar techniques

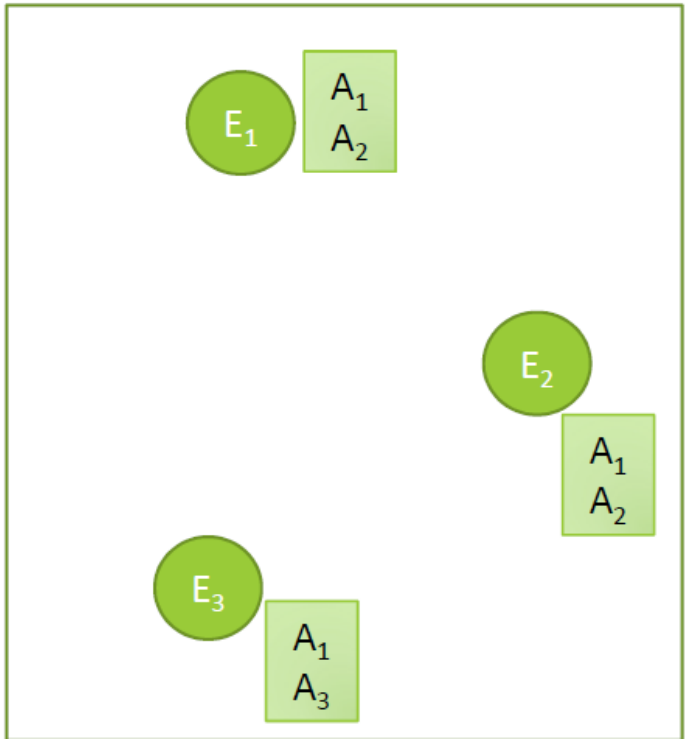
Recap: KG

- Knowledge in graph form!
- Captures entities, attributes, and relationships
- Nodes are entities



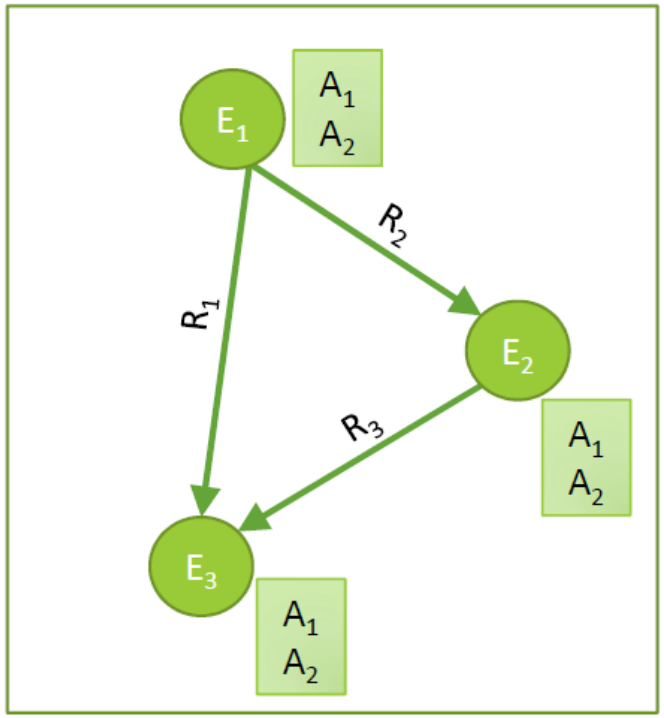
Recap: KG

- Knowledge in graph form!
- Captures entities, attributes, and relationships
- Nodes are entities
- Nodes are labeled with attributes (e.g., types)



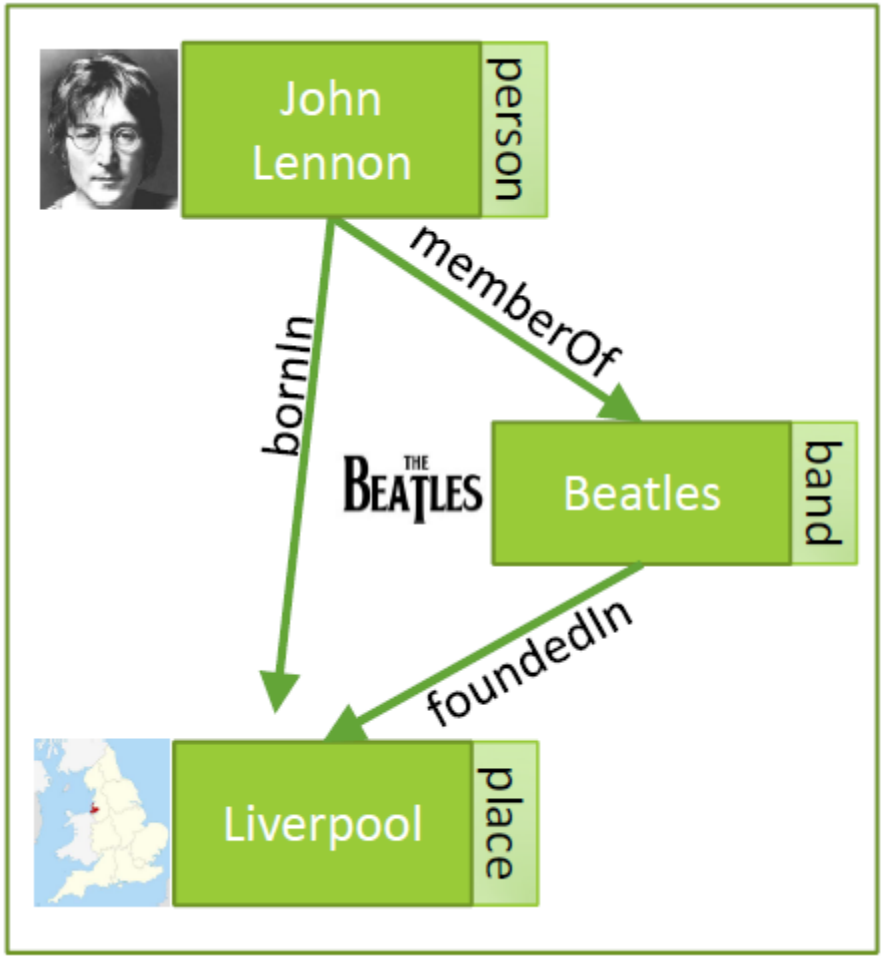
Recap: KG

- Knowledge in graph form!
- Captures entities, attributes, and relationships
- Nodes are entities
- Nodes are labeled with attributes (e.g., types)
- Typed edges between two nodes capture a relationship between entities

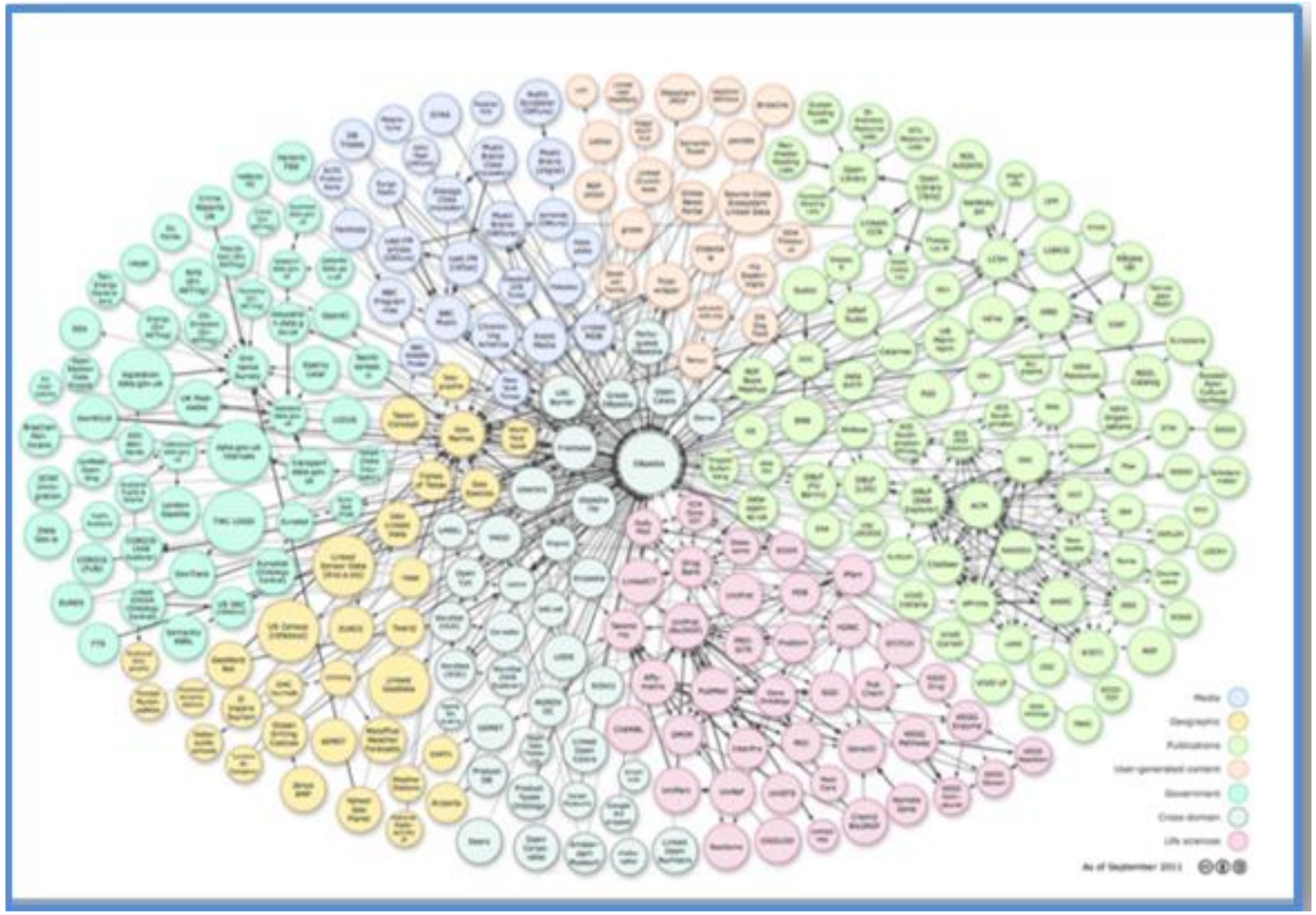


Recap: KG

- Knowledge in graph form!
- Captures entities, attributes, and relationships
- Nodes are entities
- Nodes are labeled with attributes (e.g., types)
- Typed edges between two nodes capture a relationship between entities



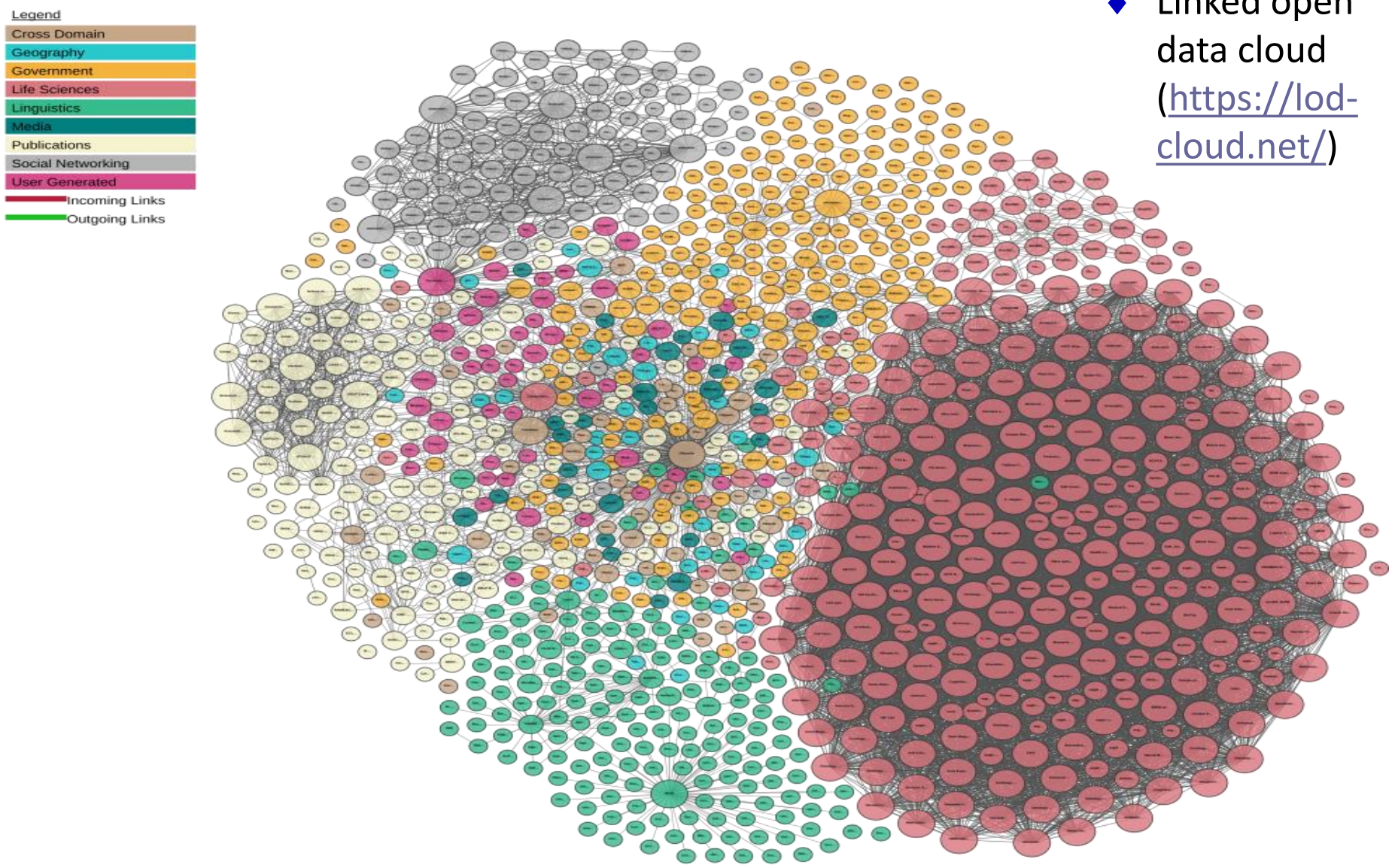
Knowledge Graphs



- Linked open data cloud (https://lod-cloud.net/)

The dataset currently contains **1,239** datasets with **16,147** links (as of March 2019)

Knowledge Graphs



- Linked open data cloud (https://lod-cloud.net/)

The dataset currently contains **1,239** datasets with **16,147** links (as of March 2019)

Knowledge Graphs

- 2007: DBpedia is released by collaborated effort of Leipzig University, University of Mannheim, and OpenLink Software; Structured information is mined from Wikipedia
- 2007: A California-based startup Metaweb releases Freebase; Structured data automatically harvested from the web
- 2008: Max-Planck-Institute in Saarbrucken releases YAGO
  - Stands for: Yet Another Great Ontology
  - Extracts information automatically from Wikipedia, WordNet, and GeoNames
- 2010: Carnegie Mellon proposes a design for a language learning system NELL
  - Stands for Never-Ending Language Learner
  - Builds a knowledge base by reading the web
- 2012: Google announces addition of knowledge graph to their search
  - Goal: make search easier for the user

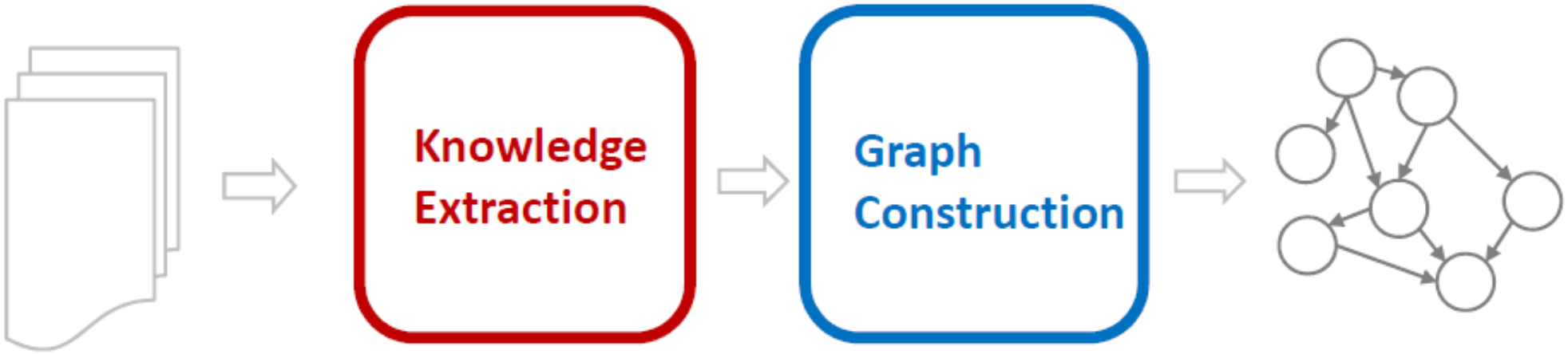
Agenda

- Recap
- Topics
- Topic allocation
- Seminar techniques

KGs: Topics

- What is a KG? -- 1
- KG construction from
  - Text --2
  - Table --3
  - Mixed --4
- KG search
  - Keyword over KG --5
  - Question answering over KG--6
  - ChatGPT for KG –7
- KG schema
  - Ontology design for KG --8
- KG alignment---9
- KG embedding ---10
- KG for ML ---11
- KG applications:
  - KG for health --- 12
  - KG for agriculture --- 13
  - KG for energy ---14
  - KG for biodiversity ---15
- KG frameworks:
  - Google KG --- 16
  - Amazon PKG --- 17
  - Dbpedia ---- 18
  - Wikidata --- 19

1. KG construction



https://kgtutorial.github.io

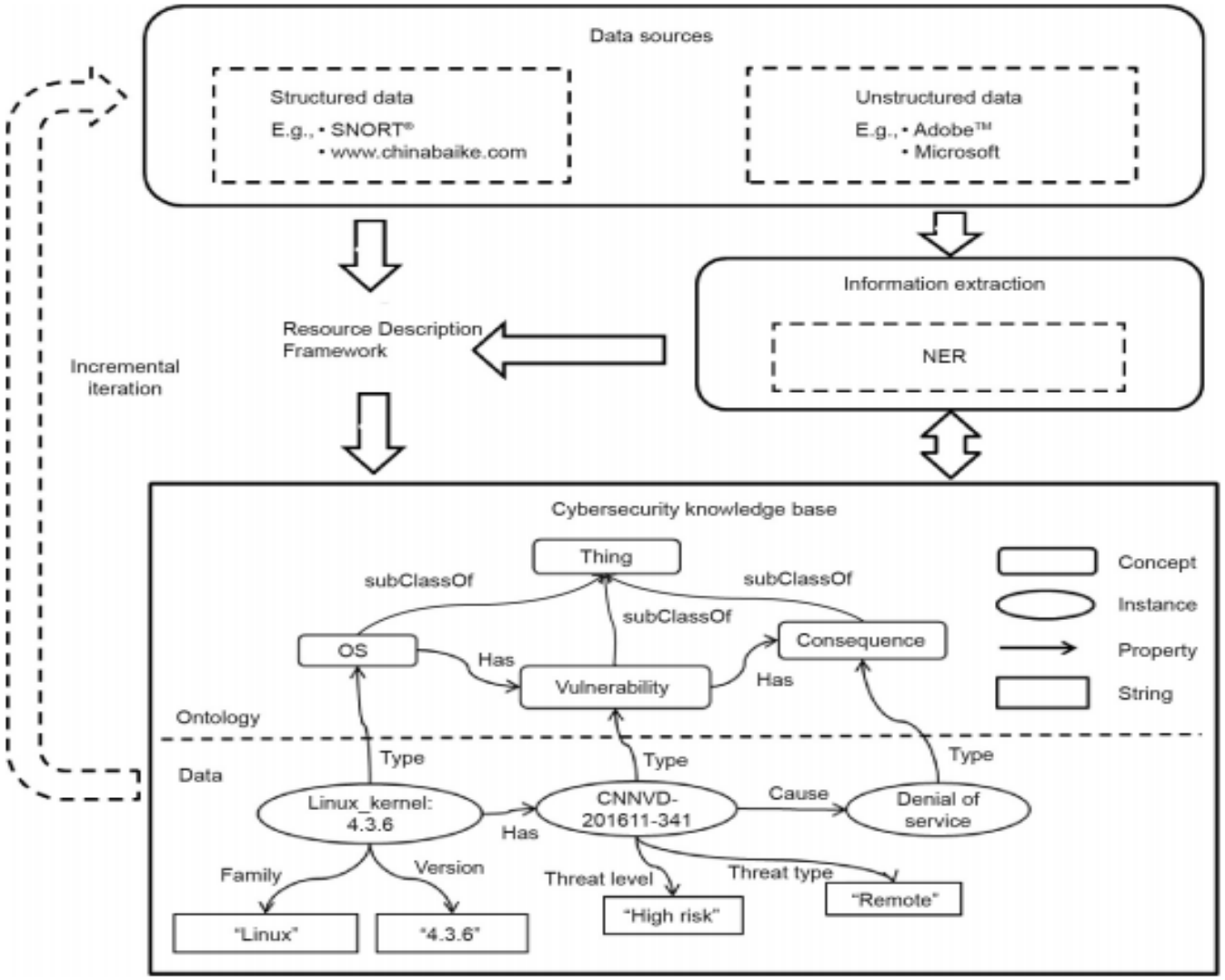
https://stiinnsbruck.github.io/kgt/

https://kdd2018tutorialt39.azurewebsites.net/

https://kg-construct.github.io/eswc-dkg-tutorial-2022/

KG Construction

- Main steps:
  - Extract
  - Representation
  - Reasoning and inference
- Who are the entities (nodes) in the graph?
- What are their attributes and types (labels)?
- How are they related (edges)?



https://www.sciencedirect.com/science/article/pii/S2095809918301097

KG Construction

John was born in Liverpool, to Julia and Alfred Lennon.

Text

NLP

Lennon..
John Lennon...

the Pool

Mrs. Lennon..
.. his mother ..

his father
he
Alfred

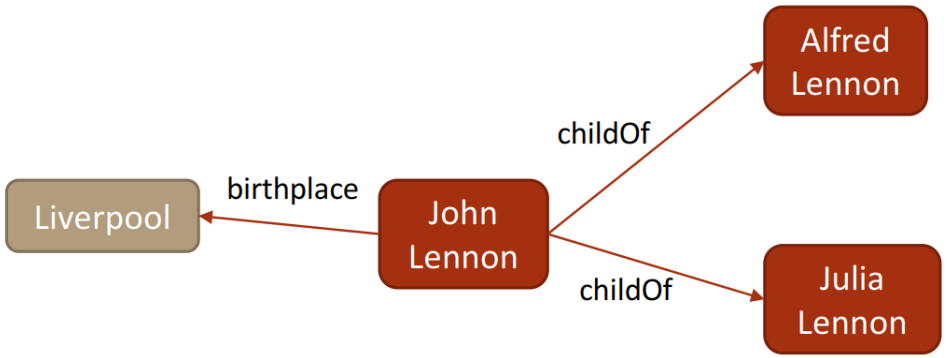
Person — Location — Person — Person

John was born in Liverpool, to Julia and Alfred Lennon.

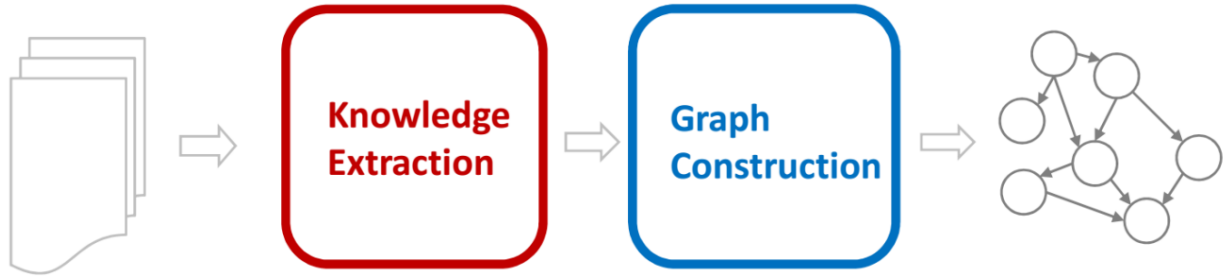
Annotated text

Information Extraction

Extraction graph

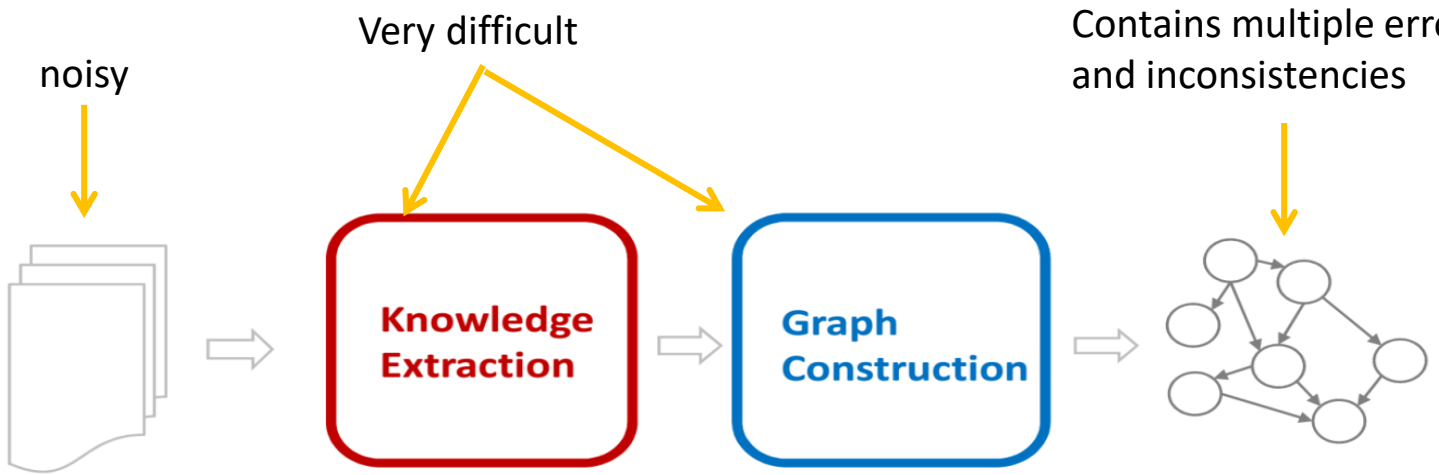


KG Construction

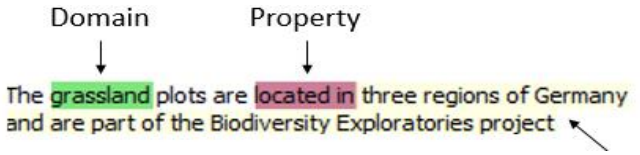


Knowledge extraction	Graph Construction
• **Who** are the entities (nodes) in the graph? • Named Entity Recognition • Entity Coreference	• **Who** are the entities (nodes) in the graph? • Entity Linking • Entity Resolution
What are their attributes and types (labels)? • Named Entity Recognition	**What** are their attributes and types (labels)? • Collective Classification
How are they related (edges)? • Relation Extraction • Semantic Role Labeling	**How** are they related (edges)? • Link Prediction

2. Quality and evaluation of KGs



- Different Web knowledge graphs have been created.
  - often constructed from semi-structured knowledge, such as Wikipedia, or harvested from the web with a combination of statistical and linguistic methods.
- The result are large-scale knowledge graphs that try to make a good trade-off between completeness and correctness.
- In order to further increase the utility of such knowledge graphs, various refinement methods have been proposed



2. Quality and evaluation of KGs

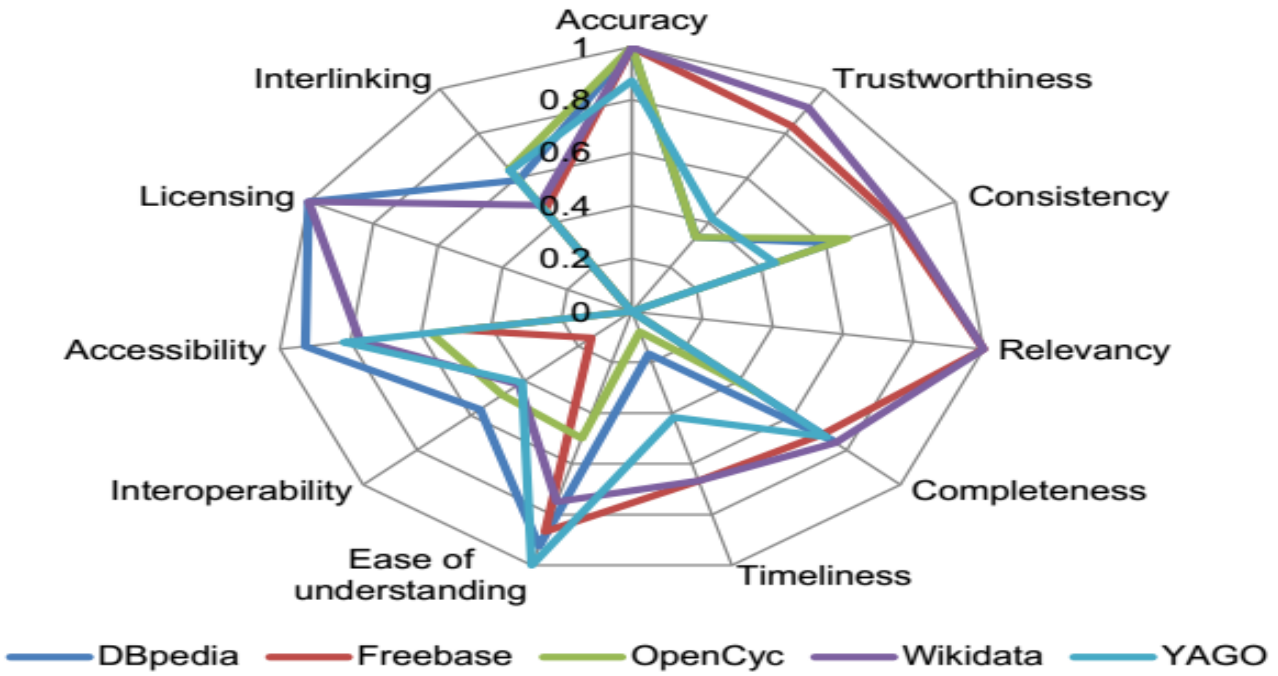
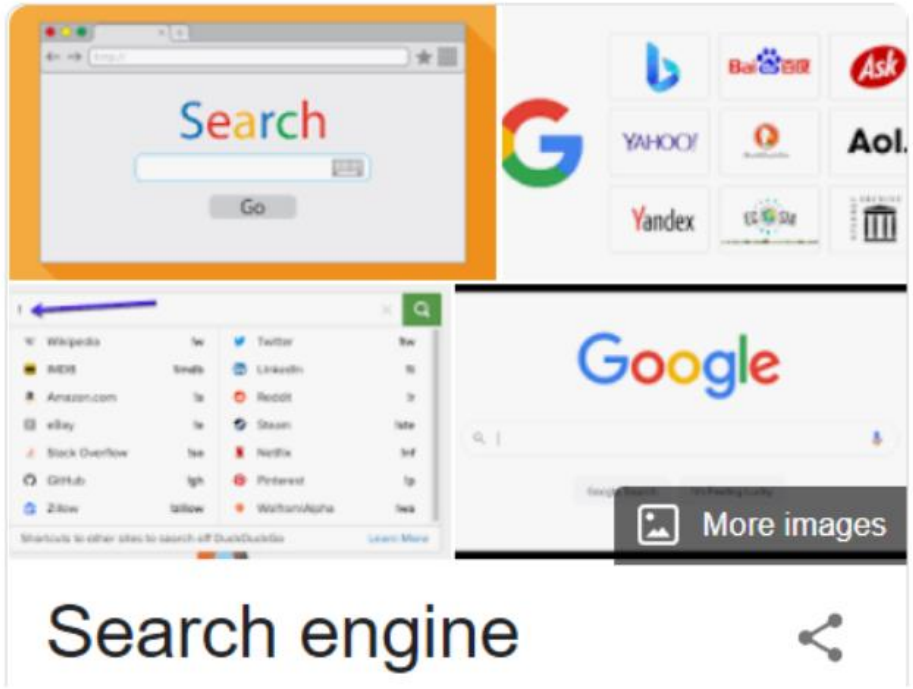


Figure 1: Results of our data quality assessment, gained by averaging the corresponding data quality metric scores for each data quality dimension.

https://arxiv.org/pdf/1809.11099.pdf

3. Search over KGs

- periodic symbol, silicon?
  - what is the periodic symbol for silicon?

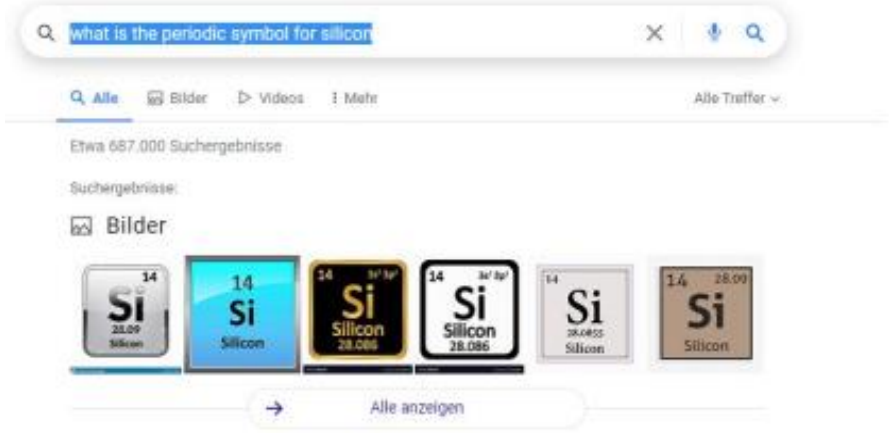


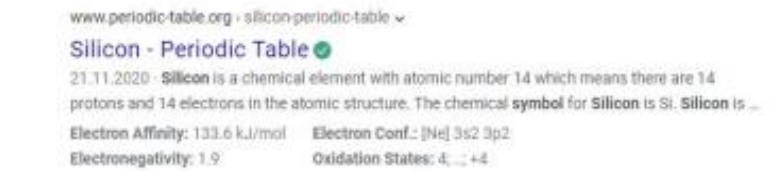
DBpedia

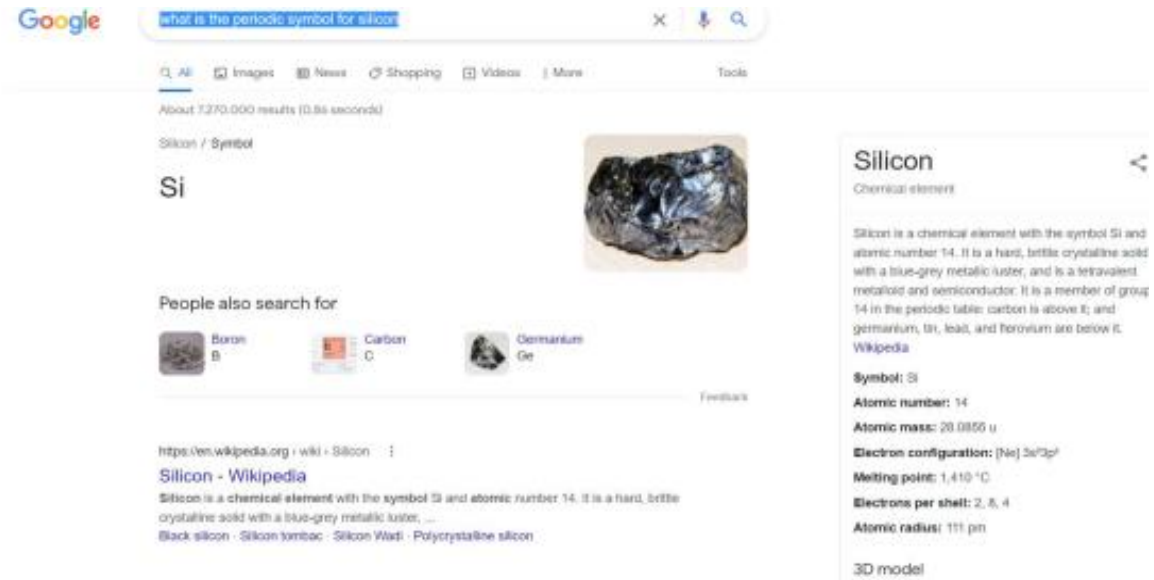
https://dbpedia.org/fct/

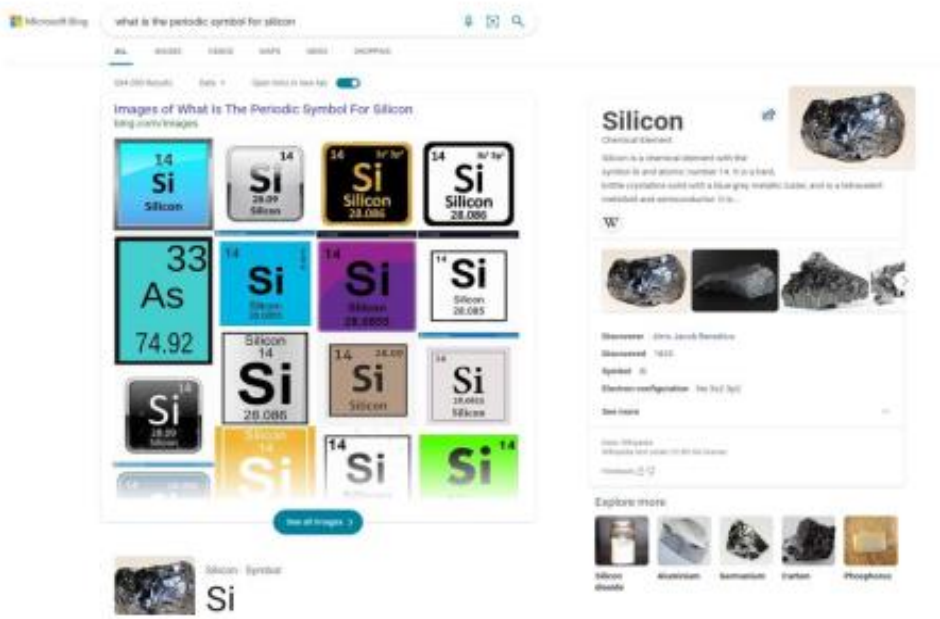
3. Search over KGs: Question answering

- what is the periodic symbol for silicon?



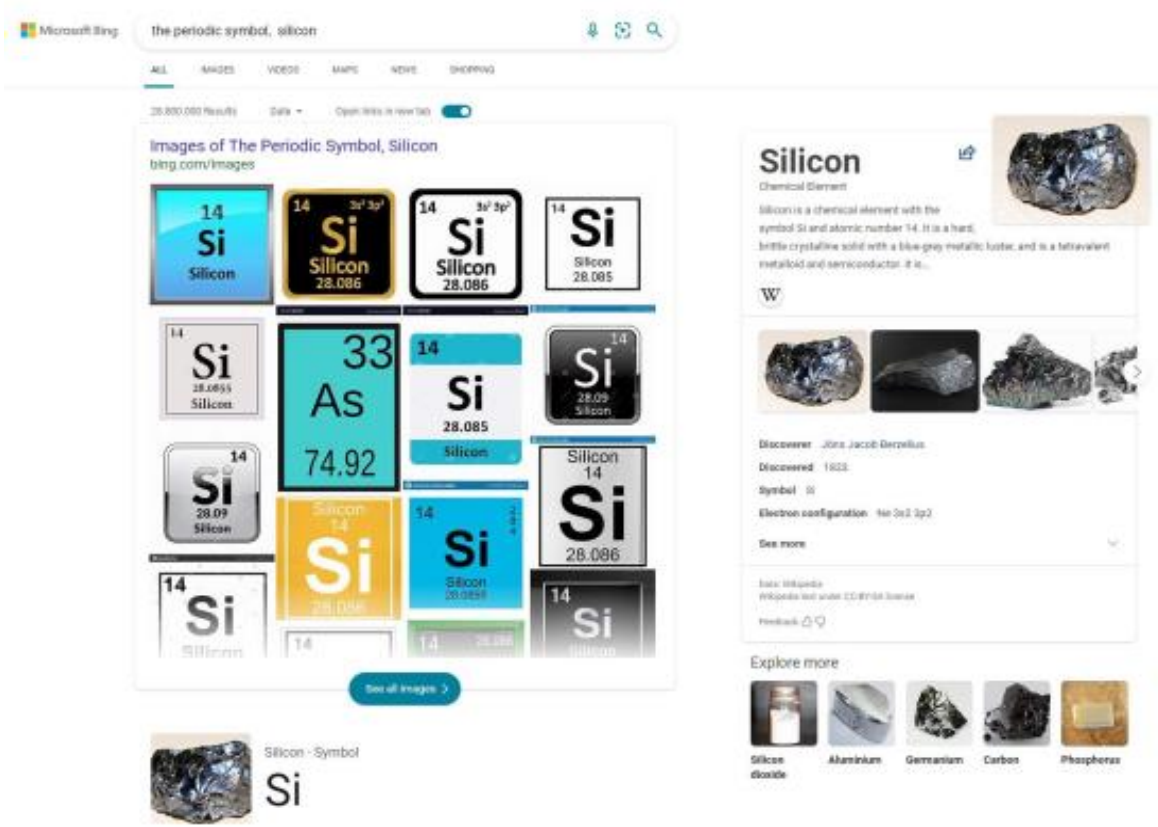


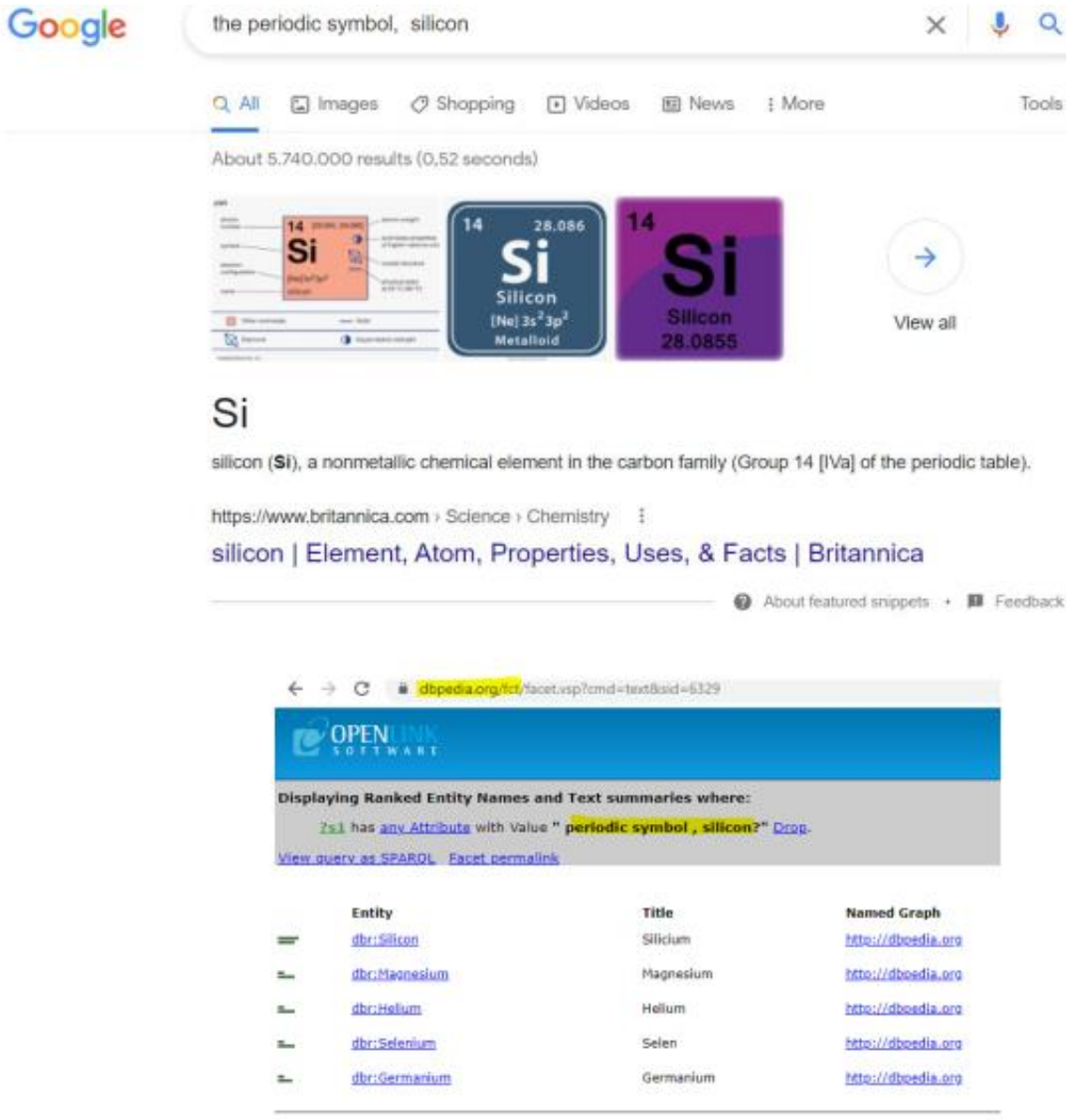




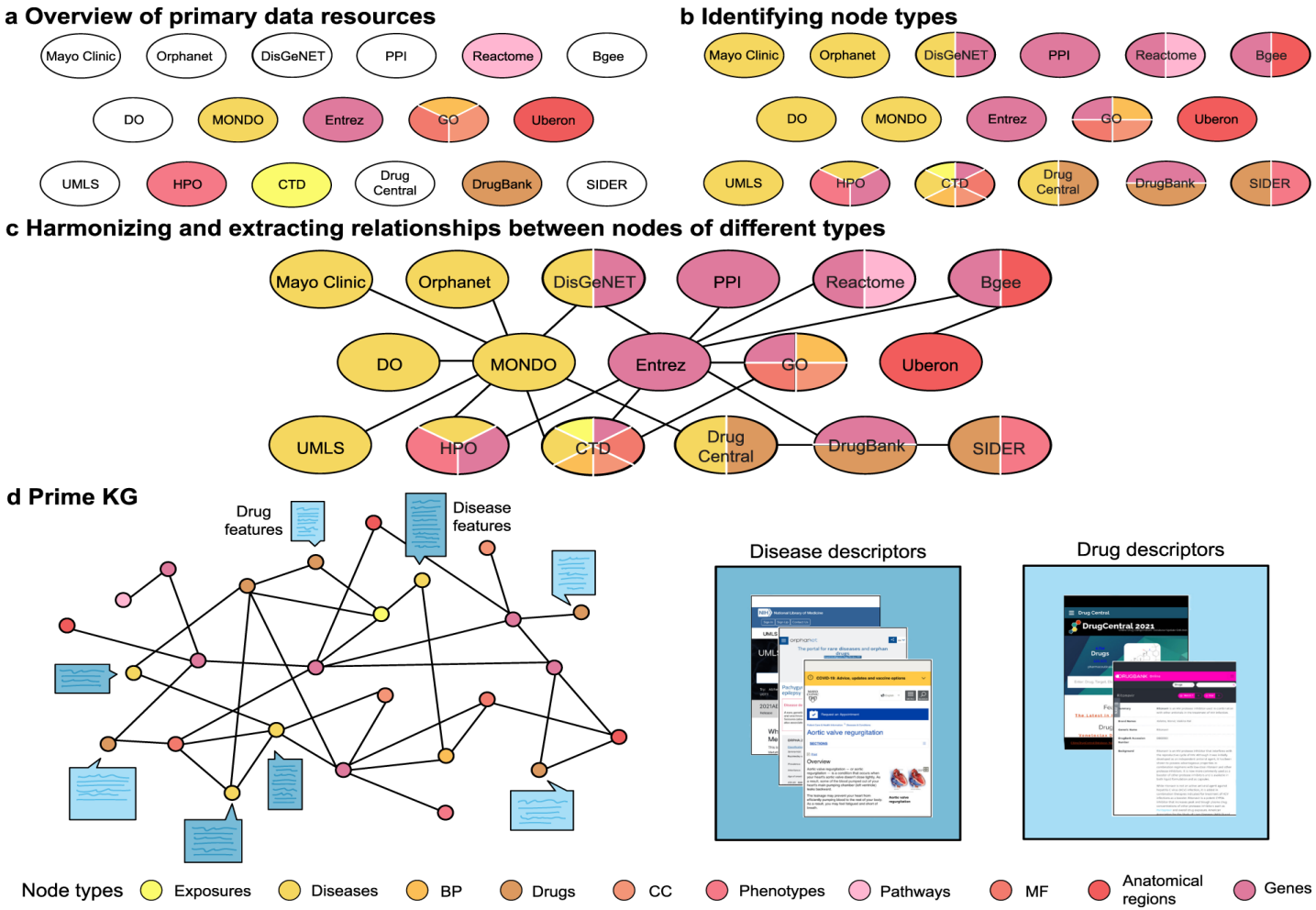
3. Search over KGs: Keyword

- periodic symbol, silicon



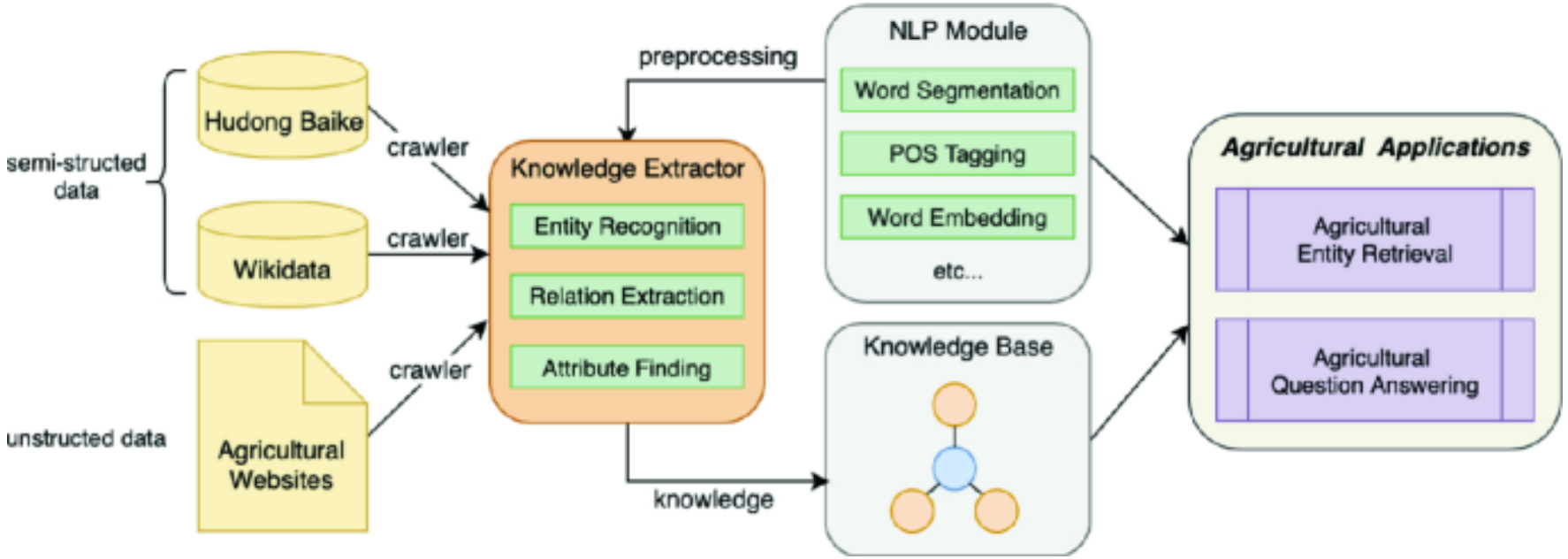


4. KGs Applications: Health



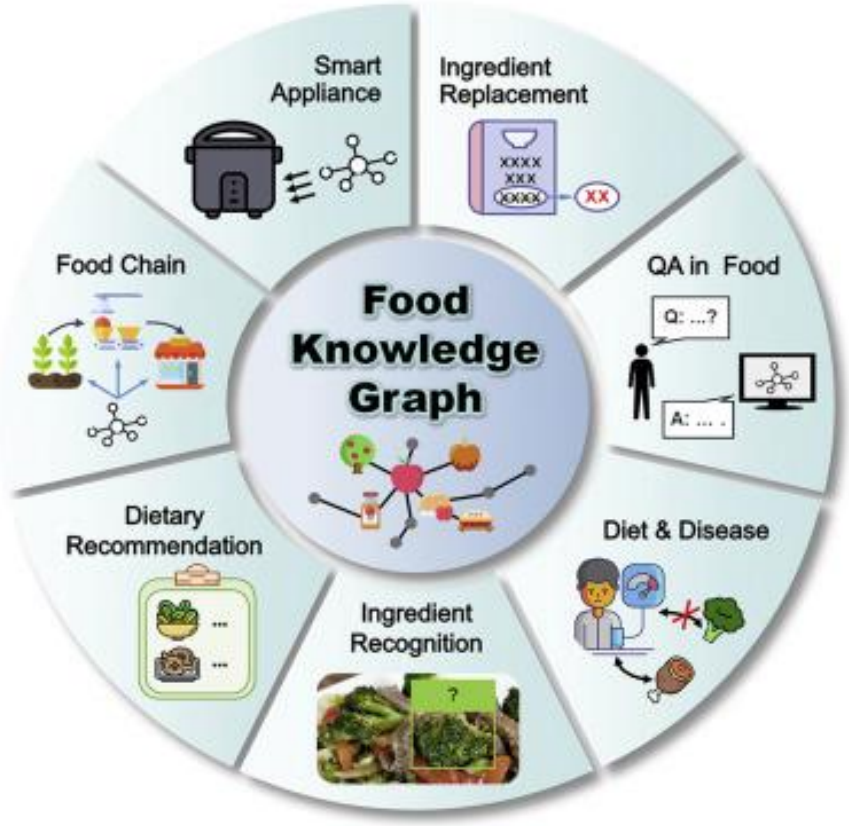
Building a knowledge graph to enable precision medicine: https://www.nature.com/articles/s41597-023-01960-3

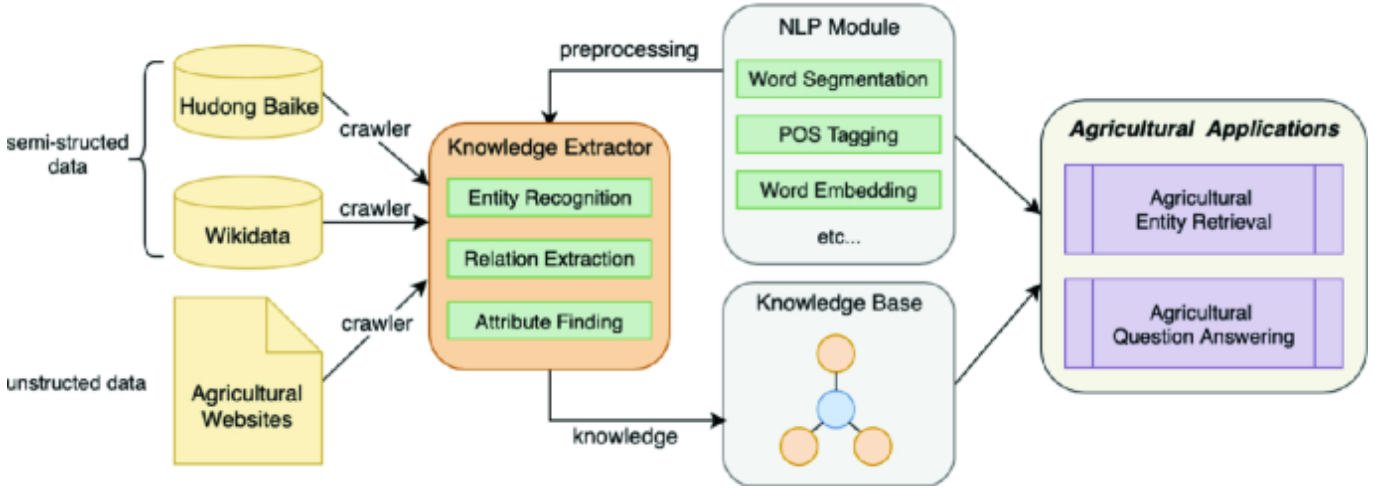
4. KGs Applications: Agriculture



AgriKG: An Agricultural Knowledge Graph and Its Applications: https://link.springer.com/chapter/10.1007/978-3-030-18590-9_81

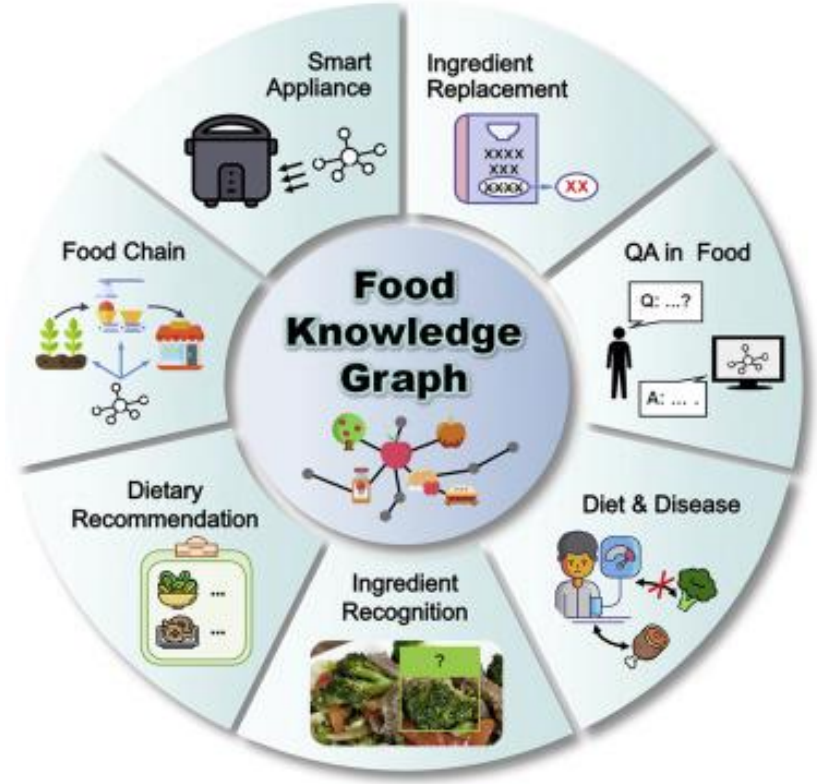
4. KGs Applications: Agriculture

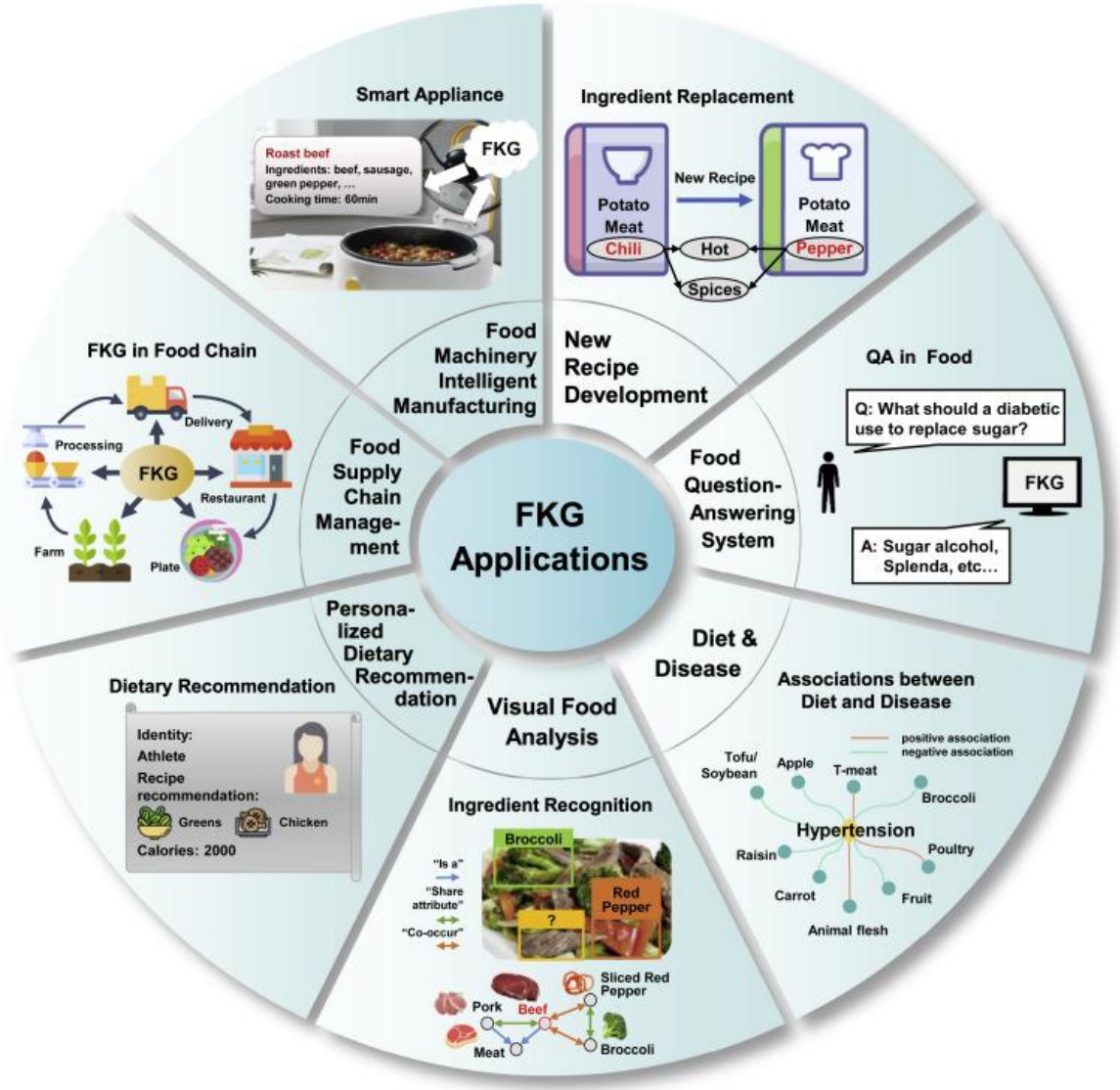




Applications of knowledge graphs for food science and industry: https://www.sciencedirect.com/science/article/pii/S2666389922000691

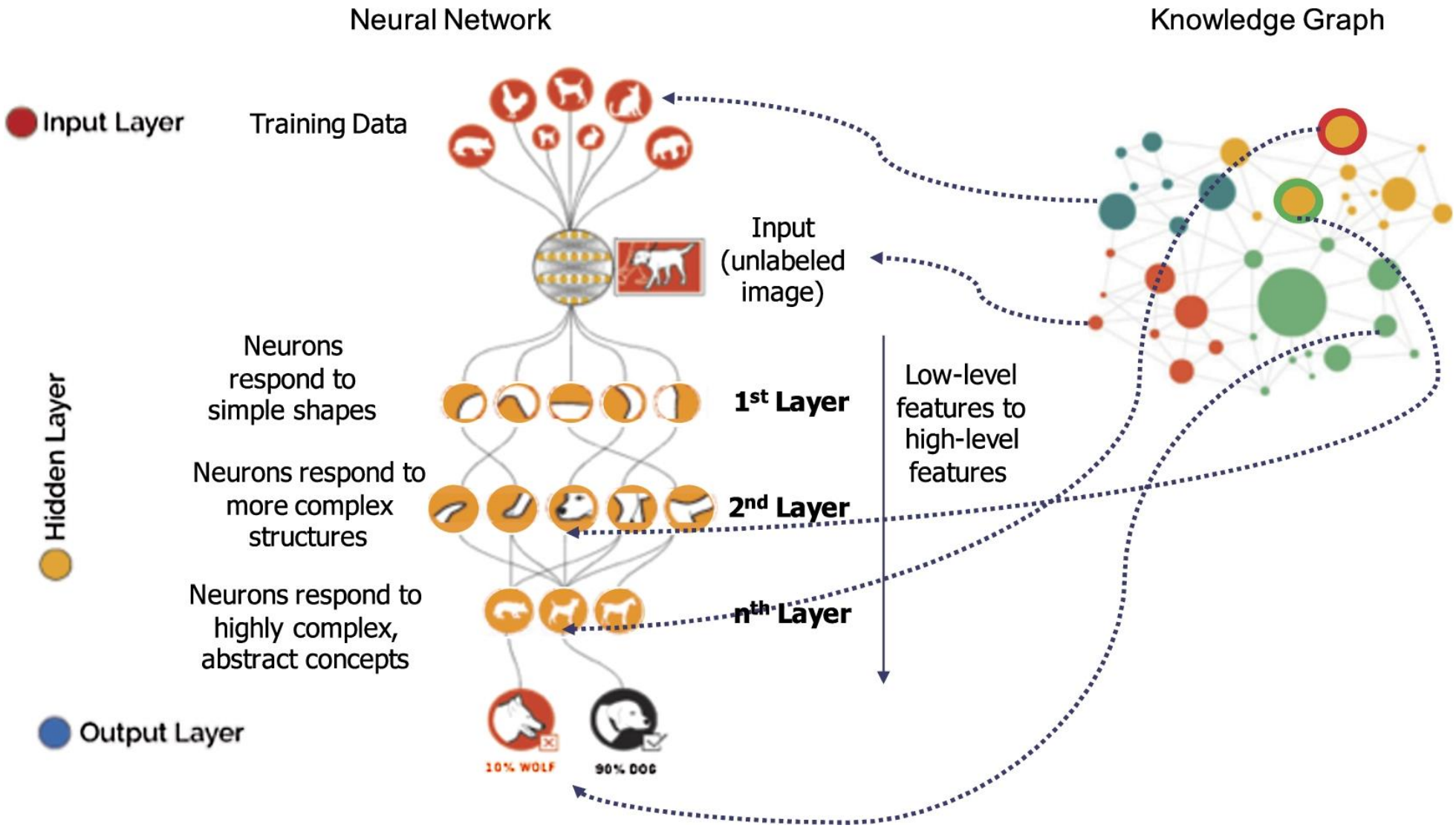
4. KGs Applications: Agriculture





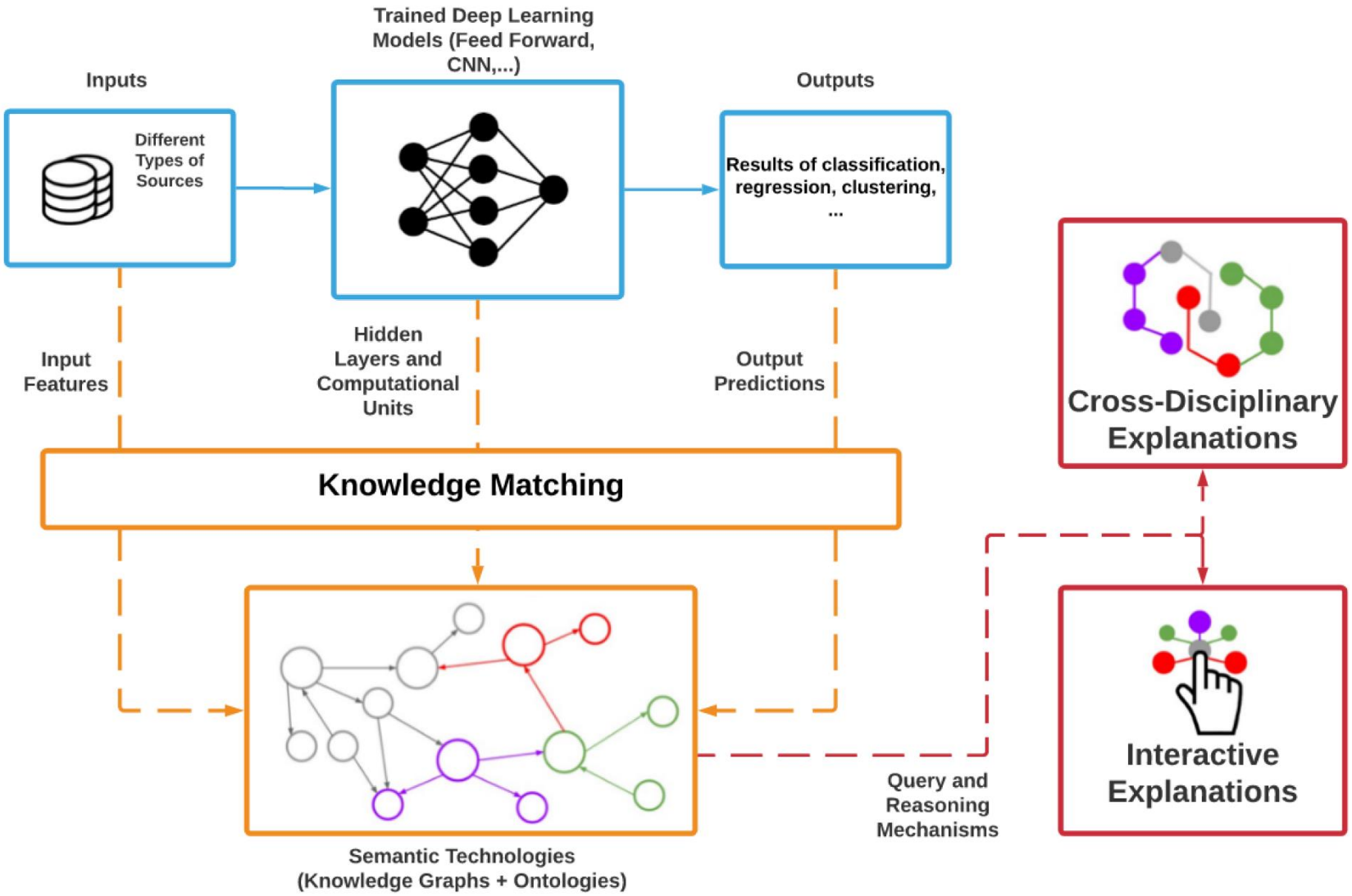
Applications of knowledge graphs for food science and industry: https://www.sciencedirect.com/science/article/pii/S2666389922000691

5. KGs for ML



Knowledge graphs as tools for explainable machine learning: A survey: https://www.sciencedirect.com/science/article/pii/S0004370221001788

5. KGs for ML

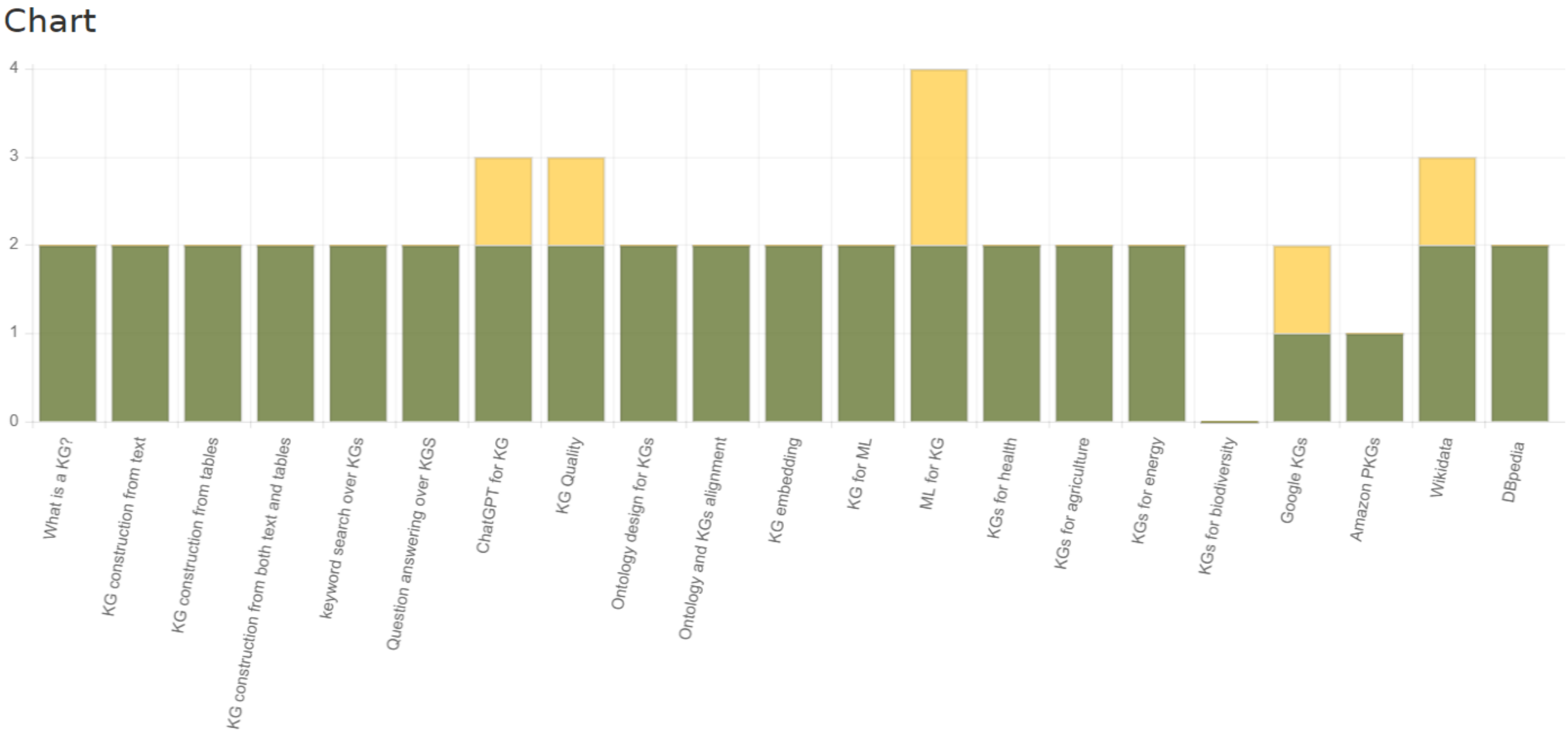


On the Integration of Knowledge Graphs into Deep Learning Models for a More Comprehensible AI—Three Challenges for Future Research

Agenda

- Recap
- Topics
- Topic allocation
- Seminar techniques

Topic allocation



Topic allocation

	What is a KG?	KG constr	KG constr	KG constr	keyword s	Question	ChatGPT f	KG Qualit	Ontology	Ontology	KG embec	KG for ML	ML for KG	KGs for he	KGs for ag	KGs for er	KGs for bi	Google KC	Amazon P	Wikidata	DBpedia
Hrishikesh Jadhav	No	No	No	No	No	No	Under res	No	No	No	No	Yes	Under res	No	No	No	No	Under res	No	No	No
Houria chiraz bou	No	Yes	No	No	No	No	No	No	No	No	No	Yes	Under res	No	No	No	No	No	No	No	No
Andreas Einwiller	No	No	No	No	No	No	No	No	No	No	Yes	No	No	No	No	No	No	No	No	No	No
Mohammadreza N	No	No	No	No	No	No	Yes	No	No	No	No	No	Yes	No	No	No	No	No	No	No	No
ilnaz tayebi	No	Yes	Yes	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No
Amandeep Singh (	No	No	No	No	Yes	No	Yes	No	No	No	No	No	No	No	No	No	No	No	No	No	No
Sarra Ben Brahim	No	No	No	No	No	No	No	No	No	No	No	No	No	Yes	No	No	No	No	No	No	No
Akanksha Vijayver	No	No	No	No	No	No	No	No	No	No	No	No	No	No	Yes	Yes	No	No	No	Yes	No
Chirag Natesh Vija	Yes	No	No	Yes	No	No	No	No	No	No	Yes	No	Yes	No	No	No	No	No	No	No	No
Vibhash Singh	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	Under res	Yes
Negin Shademan	No	No	Yes	No	Yes	Yes	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No
Elaheh Alinezhad	No	No	No	Yes	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	Yes	Yes
Anar Alimzade	No	No	No	No	No	No	No	No	No	Yes	No	No	No	No	No	No	No	No	No	No	No
Emilien Marchet	No	No	No	No	No	Yes	No	Under res	No	Yes	No	No	No	No	No	No	No	No	No	No	No
Reyhaneh Afshari	Yes	No	No	No	No	No	No	No	No	No	No	No	No	Yes	No	Yes	No	No	No	No	No
Jonas Picker	No	No	No	No	No	No	No	Yes	No	No	No	No	No	No	No	No	No	No	No	No	No
Elif GÃ¼nay	No	No	No	No	No	No	No	No	Yes	No	No	No	No	No	Yes	No	No	Yes	No	No	No
Florian RASCOUSS	No	No	No	No	No	No	No	Yes	Yes	No	No	No	No	No	No	No	No	No	No	No	No
Sami Abdel-Fattah	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	Yes	No	No

Topic allocation

topic	student	Monitor
What is a KG?	Rascoussier, Florian Guillaume Pierre	Prof. Algergawy
KG construction from text	Boudemagh, Houria-Chiraz	Prof. Algergawy
KG construction from tabular	Tayebi, Ilnaz	Prof. Algergawy
KG construction from both	Natesh Vijay, Chirag	Asha
KG quality	Picker, Jonas	Vishva
Keyword search over KG	Shademan, Negin	Vishva
QA over KG	Marchet, Emilien	Vishva
ChatGPT for KG	Gill, Amandeep Singh	Vishva
Ontology for KG	Alimzade, Anar	Vishva
KG alignment	Rastogi, Deepak	Vishva

topic	Student	Monitor
KG embedding	Einwiller, Andreas	Vishva
KG for ML	Jadhav, Hrishikesh	Asha
ML for KG	Mohebbi Najmabad, Mohammadreza	Asha
KG for agriculture	Vijayvergiya, Akanksha	Asha
KG for health	Sarra Ben Brahim	Prof. Algergawy
KG for biodiversity	Bouaoud, Saif Eddine	Prof. Algergawy
Google KG	Günay, Elif	Asha
Amazon PKG	Abdel-Fattah, Sami Raid Khalid	Asha
Wikidata	Alinezhad, Elaheh	Prof. Algergawy
DBpedia	Singh, Vibhash Kumar	Prof. Algergawy

Topic allocation

topic	Date	Moderator
What is a KG?	07.06	Einwiller, Andreas
KG construction from text		Jadhav, Hrishikesh
KG construction from tabular		Mohebbi Najmabad, Mohammadreza
KG construction from both	14.06	Abdel-Fattah, Sami Raid Khalid
KG Quality		Bouaoud, Saif Eddine
Keyword search over KG		Alinezhad, Elaheh
QA over KG	21.06	Singh, Vibhash Kumar
ChatGPT for KG		Vijayvergiya, Akanksha
Ontology for KG		Günay, Elif

topic	Date	Moderator
KG alignment	28.06	Natesh Vijay, Chirag
KG embedding		Shademan, Negin
KG for ML		Marchet, Emilien
ML for KG	05.07	Gill, Amandeep Singh
KG for agriculture		Rascoussier, Florian Guillaume Pierre
KG for biodiversity	12.07	Boudemagh, Houria-Chiraz
KG for health		Alimzade, Anar
Google KG		Tayebi, Ilnaz
Amazon PKG	19.07	Picker, Jonas
Wikidata		Sarra Ben Brahim
DBpedia		Rastogi, Deepak

Outline

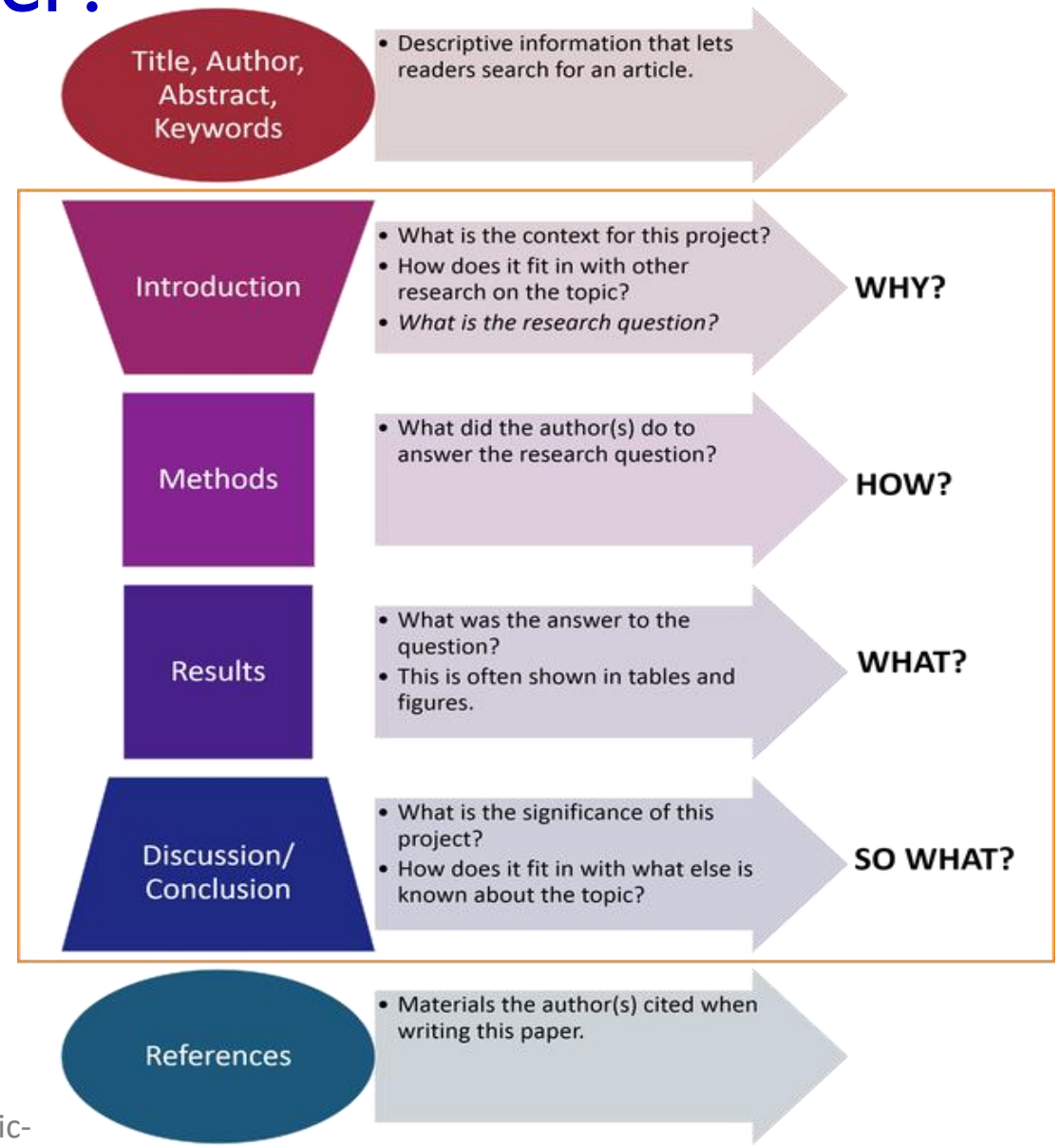
- Recap
- Topics
- Topic allocation
- Seminar techniques

Where to find research papers?

- ACM DL: https://dl.acm.org/
- DBLP: https://dblp.org/
- IEEE explorer: https://ieeexplore.ieee.org/Xplore/home.jsp
- Google scholar: https://scholar.google.com/

- Others:
  - ScienceDirect: - https://www.sciencedirect.com/
  - CiteSeerx: https://citeseerx.ist.psu.edu/
  - semantic scholar: https://www.semanticscholar.org/
  - ArXiv: https://arxiv.org/

How can I read a scientific paper?



https://library.sumdu.edu.ua/en/for-researcher/academic-writing/writing-scientific-text/the-structure-of-scientific-texts/823-the-structure-of-scientific-texts.html

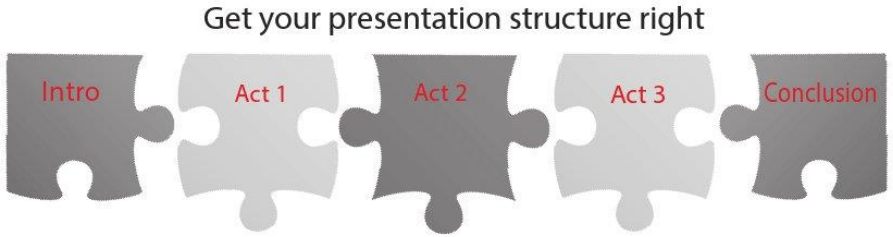
How can I read a scientific paper?

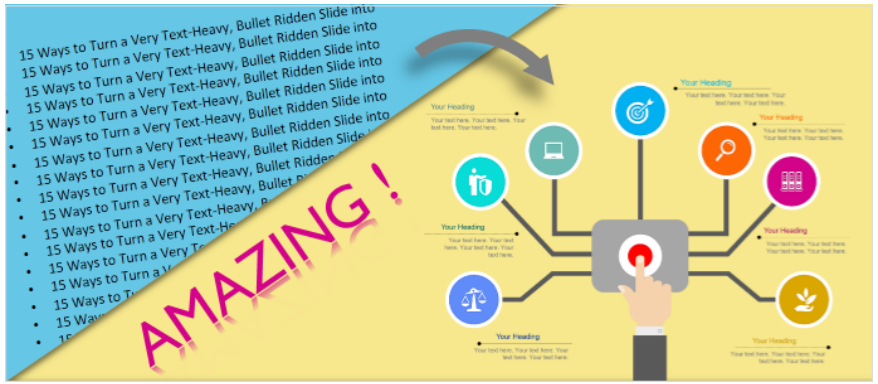
- Step 0: start by asking yourself the following question: “Why am I reading this paper?”
- Step 1: Skim the article.: Summarize the all article in 5-10 lines giving anticipation of the findings
- Step 2. Grasp the vocabulary. Begin to go through the article and highlight words and phrases you do not understand
- Step 3. Identify the structure of the article and work on your comprehension
- Step 4. Read the bibliography/references section. Reading the references or works cited may lead you to other useful resources
- Step 5. Reflect on what you have read and draw your own conclusions
- Step 6. Read the article a second time in chronological order. Reading the article a second time will reinforce your overall understanding.

https://francescolelli.info/thesis/read-scientific-papers-quickly-and-effectively/

Why my presentation is too boring?

- Try to avoid:
  - No presentation structure
  - Too much content
  - Visually unappealing
  - Unengaging content





How to prepare a good scientific report?

- "Scientific writing is very precise, so it's important to make sure you're as concise and clear as possible. Being clear with your purpose helps you stay focused on what you're writing about"
  — Dennis Farrugia, Language and Learning Adviser

- Report structure
  - *Title:* Use a descriptive and meaningful title
  - *Abstract*: first contact between the report and the reader, where you summarise what you did, how you did it and your results
  - *Introduction*: a more detailed abstract and an introduction to the topic in general
  - *Main text*: describe the actual problem, solution, main methods, include figures, equations
  - *Results*:
  - Summary/conclusions
  - References

https://vcai.mpi-inf.mpg.de/teaching/how_to_thesis/how_to_write_a_report_slussalek.pdf

Grading

- Presentation
  - Structure and organization
  - Visually appealing
  - Engaging content
  - Q/A section

- Report
  - Structure and organization
  - Correctness of content, thematic completeness with respect to available space and understandability
  - Plagiarism check


QUESTIONS
ANSWERS