

Knowledge Graph Construction, Evaluation, and Applications

Haihua Chen

March 1, 2022



Outline

- ▶ **Introduction to Knowledge Graph**

Outline

- ▶ **Introduction to Knowledge Graph**
- ▶ **How to Construct a Large-scale Knowledge Graph**

Outline

- ▶ **Introduction to Knowledge Graph**
- ▶ **How to Construct a Large-scale Knowledge Graph**
- ▶ **Towards a Legal Knowledge Graph**

Outline

- ▶ **Introduction to Knowledge Graph**
- ▶ **How to Construct a Large-scale Knowledge Graph**
- ▶ **Towards a Legal Knowledge Graph**
- ▶ **Knowledge Graph Qualify and Evaluation**

Outline

- ▶ **Introduction to Knowledge Graph**
- ▶ **How to Construct a Large-scale Knowledge Graph**
- ▶ **Towards a Legal Knowledge Graph**
- ▶ **Knowledge Graph Qualify and Evaluation**
- ▶ **Knowledge Graph Applications**

Introduction to KG

What is a Knowledge Graph (KG)

- ▶ Knowledge graph (KG) is a multi-relational graph composed of entities and relations which are regarded as nodes and different types of edges, respectively ¹.

¹ S. Ji, et al. "A Survey on Knowledge Graphs: Representation, Acquisition and Applications." arXiv preprint arXiv:2002.00388 (2020).

Introduction to KG

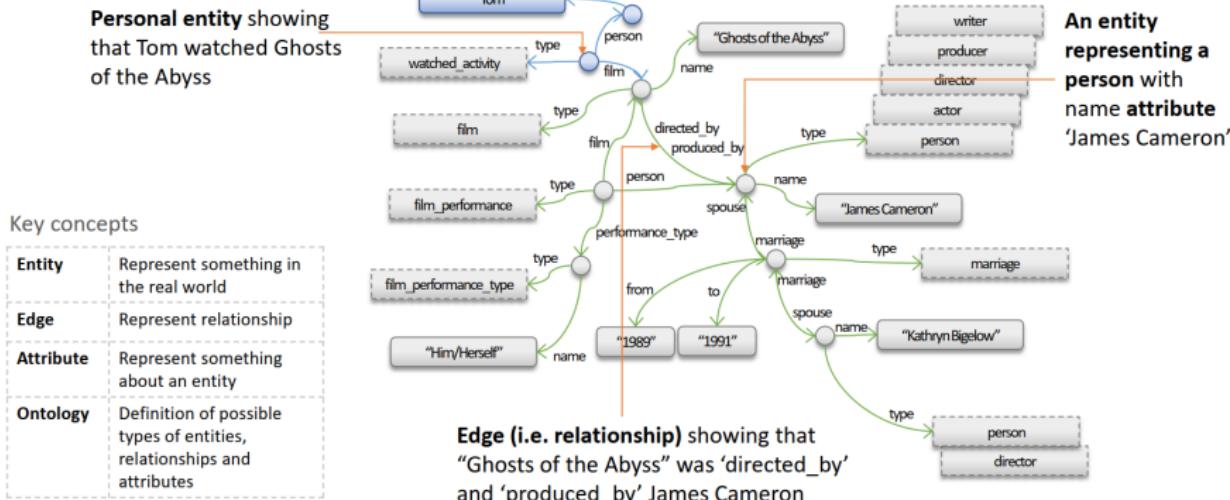
What is a Knowledge Graph (KG)

- ▶ Knowledge graph (KG) is a multi-relational graph composed of entities and relations which are regarded as nodes and different types of edges, respectively ¹.
- ▶ KG Representation: $\langle head, relationship, tail \rangle$.

¹S. Ji, et al. "A Survey on Knowledge Graphs: Representation, Acquisition and Applications." arXiv preprint arXiv:2002.00388 (2020).

Introduction to KG

A KG Example in the Film Domain²



²Y. Gao, et al. "Building a large-scale, accurate and fresh knowledge graph." KDD-2018, Tutorial 39 (2018).

Introduction to KG

Node

- ▶ **Entity (Object/Instance):** Something that exists as itself, as a subject or an object, actually or potentially, connectedly, or abstractly, physically or not. (Wikipedia).

Introduction to KG

Node

- ▶ **Entity (Object/Instance):** Something that exists as itself, as a subject or an object, actually or potentially, connectedly, or abstractly, physically or not. (Wikipedia).
- ▶ **Concept:** A fundamental category of existence; Representations of categories.

Introduction to KG

Node

- ▶ **Entity (Object/Instance):** Something that exists as itself, as a subject or an object, actually or potentially, connectedly, or abstractly, physically or not. (Wikipedia).
- ▶ **Concept:** A fundamental category of existence; Representations of categories.
- ▶ **Category(Type, Class):** Grouping of entities that have something in common.

Introduction to KG

Node

- ▶ **Entity (Object/Instance):** Something that exists as itself, as a subject or an object, actually or potentially, connectedly, or abstractly, physically or not. (Wikipedia).
- ▶ **Concept:** A fundamental category of existence; Representations of categories.
- ▶ **Category(Type, Class):** Grouping of entities that have something in common.
- ▶ **Value:** Data, String, Numerical, ect.

Introduction to KG

Node

- ▶ **Entity (Object/Instance):** Something that exists as itself, as a subject or an object, actually or potentially, connectedly, or abstractly, physically or not. (Wikipedia).
- ▶ **Concept:** A fundamental category of existence; Representations of categories.
- ▶ **Category(Type, Class):** Grouping of entities that have something in common.
- ▶ **Value:** Data, String, Numerical, ect.

Introduction to KG

Node

- ▶ **Entity (Object/Instance):** Something that exists as itself, as a subject or an object, actually or potentially, connectedly, or abstractly, physically or not. (Wikipedia).
- ▶ **Concept:** A fundamental category of existence; Representations of categories.
- ▶ **Category(Type, Class):** Grouping of entities that have something in common.
- ▶ **Value:** Data, String, Numerical, ect.

Introduction to KG

Edge

- ▶ **Relationship:** The relation between entities.

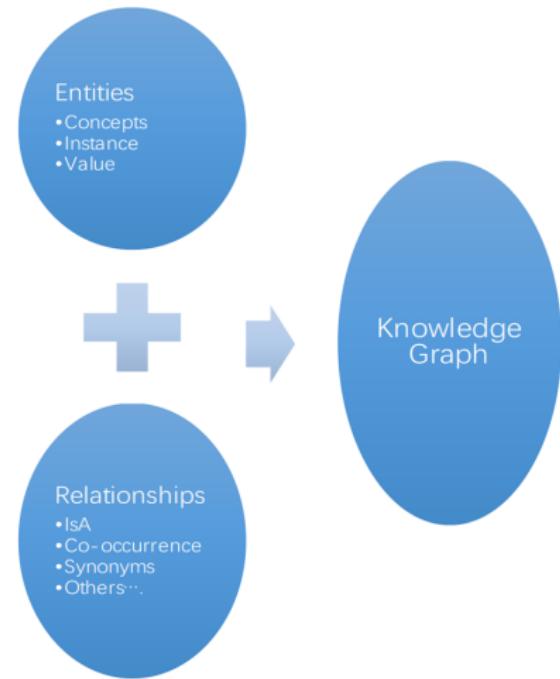
Introduction to KG

Edge

- ▶ **Relationship:** The relation between entities.
- ▶ **Attribute/Property/Quality:** A characteristic/quality to define an entity.

Introduction to KG

Model of KG³



³Y. Xiao. "Introduction to Knowledge Graphs". Fudan University, 2017.

Introduction to KG

Knowledge Graphs in the Wild ⁴

- ▶ Building knowledge graphs are expensive.

⁴ Elias and Umutcan. "Building a large-scale, accurate and fresh knowledge graph." Semantics Conference 2019, Tutorial.

Introduction to KG

Knowledge Graphs in the Wild ⁴

- ▶ Building knowledge graphs are expensive.
- ▶ Scaling them is challenging.

⁴ Elias and Umutcan. "Building a large-scale, accurate and fresh knowledge graph." Semantics Conference 2019, Tutorial.

Introduction to KG

Knowledge Graphs in the Wild ⁴

- ▶ Building knowledge graphs are expensive.
- ▶ Scaling them is challenging.
- ▶ The quality of KGs is difficult to be guaranteed and evaluated.

⁴ Elias and Umutcan. "Building a large-scale, accurate and fresh knowledge graph." Semantics Conference 2019, Tutorial.

Introduction to KG

Knowledge Graphs in the Wild⁴

- ▶ Building knowledge graphs are expensive.
- ▶ Scaling them is challenging.
- ▶ The quality of KGs is difficult to be guaranteed and evaluated.
- ▶ Knowledge Graphs:

| Name | Instances | Facts | Types | Relations |
|--------------------------|-------------|----------------|---------|-----------|
| DBpedia (English) | 4,806,150 | 176,043,129 | 735 | 2,813 |
| YAGO | 4,595,906 | 25,946,870 | 488,469 | 77 |
| Freebase | 49,947,845 | 3,041,722,635 | 26,507 | 37,781 |
| Wikidata | 15,602,060 | 65,993,797 | 23,157 | 1,673 |
| NELL | 2,006,896 | 432,845 | 285 | 425 |
| OpenCyc | 118,499 | 2,413,894 | 45,153 | 18,526 |
| Google's Knowledge Graph | 570,000,000 | 18,000,000,000 | 1,500 | 35,000 |
| Google's Knowledge Vault | 45,000,000 | 271,000,000 | 1,100 | 4,469 |
| Yahoo! Knowledge Graph | 3,443,743 | 1,391,054,990 | 250 | 800 |

⁴ Elias and Umutcan. "Building a large-scale, accurate and fresh knowledge graph." Semantics Conference 2019, Tutorial.

Introduction to KG

Common Characteristics of KGs⁵

| | Data model | Size of the graph | Development stage |
|-----------|---|---|--|
| Microsoft | The types of entities, relations, and attributes in the graph are defined in an ontology. | ~2 billion primary entities, ~55 billion facts | Actively used in products |
| Google | Strongly typed entities, relations with domain and range inference | 1 billion entities, 70 billion assertions | Actively used in products |
| Facebook | All of the attributes and relations are structured and strongly typed, and optionally indexed to enable efficient retrieval, search, and traversal. | ~50 million primary entities, ~500 million assertions | Actively used in products |
| eBay | Entities and relation, well-structured and strongly typed | Expect around 100 million products, >1 billion triples | Early stages of development and deployment |
| IBM | Entities and relations with evidence information associated with them. | Various sizes. Proven on scales documents >100 million, relationships >5 billion, entities >100 million | Actively used in products and by clients |

⁵ <https://queue.acm.org/detail.cfm?id=3332266>

Introduction to KG

The Advantages of KG

- ▶ High coverage over entities and concepts.

Introduction to KG

The Advantages of KG

- ▶ High coverage over entities and concepts.
- ▶ High coverage over semantics relationships.

Introduction to KG

The Advantages of KG

- ▶ High coverage over entities and concepts.
- ▶ High coverage over semantics relationships.
- ▶ Structured Organization.

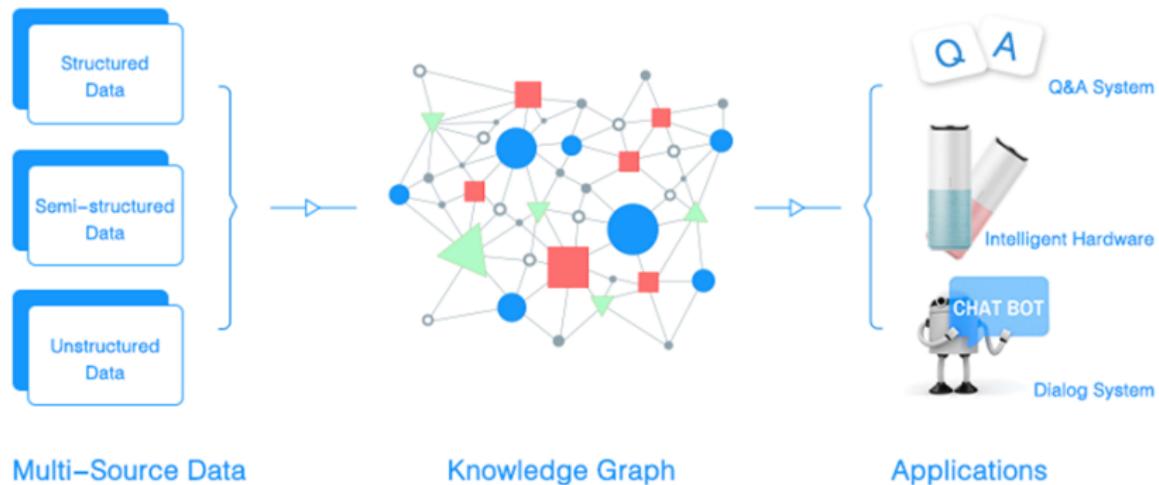
Introduction to KG

The Advantages of KG

- ▶ High coverage over entities and concepts.
- ▶ High coverage over semantics relationships.
- ▶ Structured Organization.
- ▶ Provide users with explainable results.

Introduction to KG

From KG Construction to KG Applications



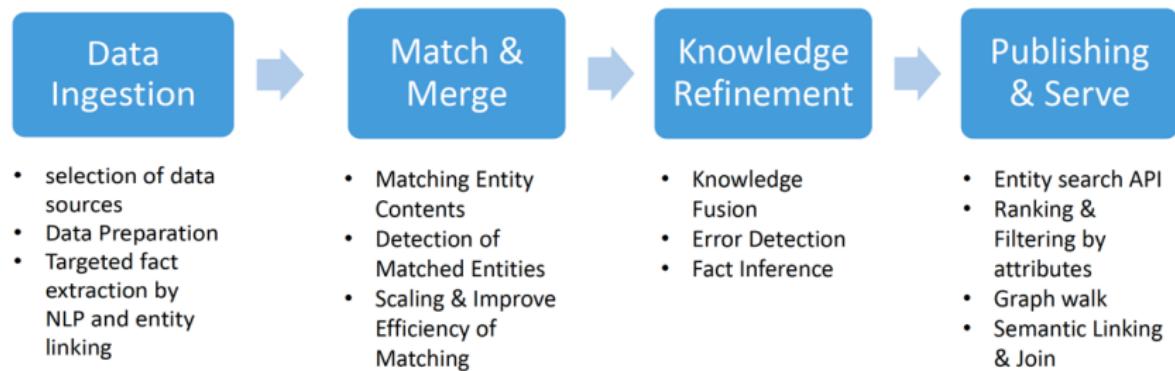
Multi-Source Data

Knowledge Graph

Applications

Construction of large-scale KG

The Pipeline of KG Construction ⁶



⁶Y. Gao, et al. "Building a large-scale, accurate and fresh knowledge graph." KDD-2018, Tutorial 39 (2018).

Construction of large-scale KG

Data Preparation ⁷

- ▶ Storing the data in a uniform manner.
 - ▶ Parsing: locate, identify, and separate data items.
 - ▶ Data transformation and standardization.

⁷Y. Gao, et al. "Building a large-scale, accurate and fresh knowledge graph." KDD-2018, Tutorial 39 (2018). 30 / 85

Construction of large-scale KG

Data Preparation ⁷

- ▶ Storing the data in a uniform manner.
 - ▶ Parsing: locate, identify, and separate data items.
 - ▶ Data transformation and standardization.
- ▶ Identify which fields to be compared.

⁷Y. Gao, et al. "Building a large-scale, accurate and fresh knowledge graph." KDD-2018, Tutorial 39 (2018). 31 / 85

Construction of large-scale KG

Data Preparation ⁷

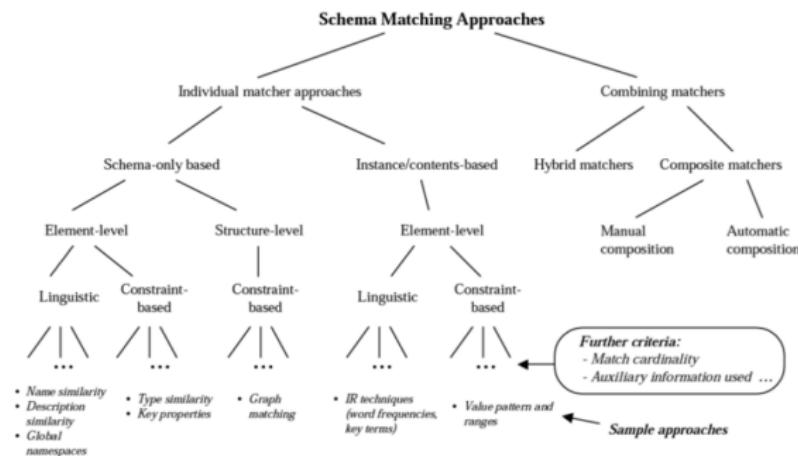
- ▶ Storing the data in a uniform manner.
 - ▶ Parsing: locate, identify, and separate data items.
 - ▶ Data transformation and standardization.
- ▶ Identify which fields to be compared.
- ▶ Schema matching.

⁷Y. Gao, et al. "Building a large-scale, accurate and fresh knowledge graph." KDD-2018, Tutorial 39 (2018). 32 / 85

Construction of large-scale KG

Data Preparation ⁷

- ▶ Storing the data in a uniform manner.
 - ▶ Parsing: locate, identify, and separate data items.
 - ▶ Data transformation and standardization.
- ▶ Identify which fields to be compared.
- ▶ Schema matching.

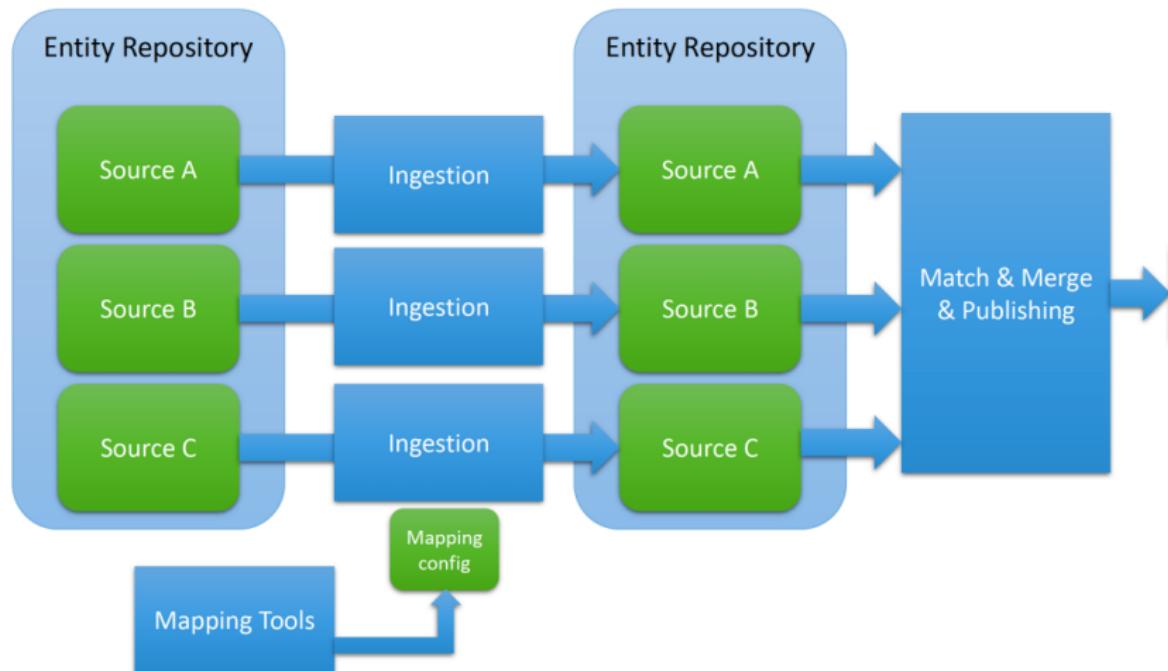


⁷

Y. Gao, et al. "Building a large-scale, accurate and fresh knowledge graph." KDD-2018, Tutorial 39 (2018).

Construction of large-scale KG

Ingestion Flow⁸



⁸

Y. Gao, et al. "Building a large-scale, accurate and fresh knowledge graph." KDD-2018, Tutorial 39 (2018).

Construction of large-scale KG

Entity Mapping⁹

Entity Mapping: To identify and discover instances referring to the same real-world entity.

► **Objective:**

- Data Enrichment.
- Improve Data Quality by identifying and removing duplicates.
- Supporting fact correctness by merging duplicate facts from multiple sources.

⁹Y. Gao, et al. "Building a large-scale, accurate and fresh knowledge graph." KDD-2018, Tutorial 39 (2018).

Construction of large-scale KG

Entity Mapping⁹

Entity Mapping: To identify and discover instances referring to the same real-world entity.

▶ **Objective:**

- ▶ Data Enrichment.
- ▶ Improve Data Quality by identifying and removing duplicates.
- ▶ Supporting fact correctness by merging duplicate facts from multiple sources.

▶ **Synonyms:** Entity Linking, Entity Resolution, Reference Reconciliation, Deduplication, Match/Merge, Merge/Purge.

⁹Y. Gao, et al. "Building a large-scale, accurate and fresh knowledge graph." KDD-2018, Tutorial 39 (2018).

Construction of large-scale KG

Knowledge Fusion (Merging Entities) ¹⁰

After merging entity nodes in the graph, we end up with conflicting facts and connections.

► Resolving facts (and finding truth)

- Majority Voting
- Identify Authoritative Sources
- Fact Checker
- Gather evidence from different sources
- Evaluate evidences
- Model joint interactions
- Aggregate evidence and predict

¹⁰Y. Gao, et al. "Building a large-scale, accurate and fresh knowledge graph." KDD-2018, Tutorial 39 (2018).

Construction of large-scale KG

Error Detection¹¹

- ▶ Data Quality Rules
 - ▶ Functional Dependency and its conditional variation, e.g.;
 $Zip \rightarrow City$
 - ▶ Inconsistency
 - Entity cannot be a movie and book
 - Date-of-birth < date-of-death
- ▶ Outliers detection
- ▶ External signals for relationship validation (e.g.; co-clicks)
- ▶ NLP features

¹¹Y. Gao, et al. "Building a large-scale, accurate and fresh knowledge graph." KDD-2018, Tutorial 39 (2018).

Construction of large-scale KG

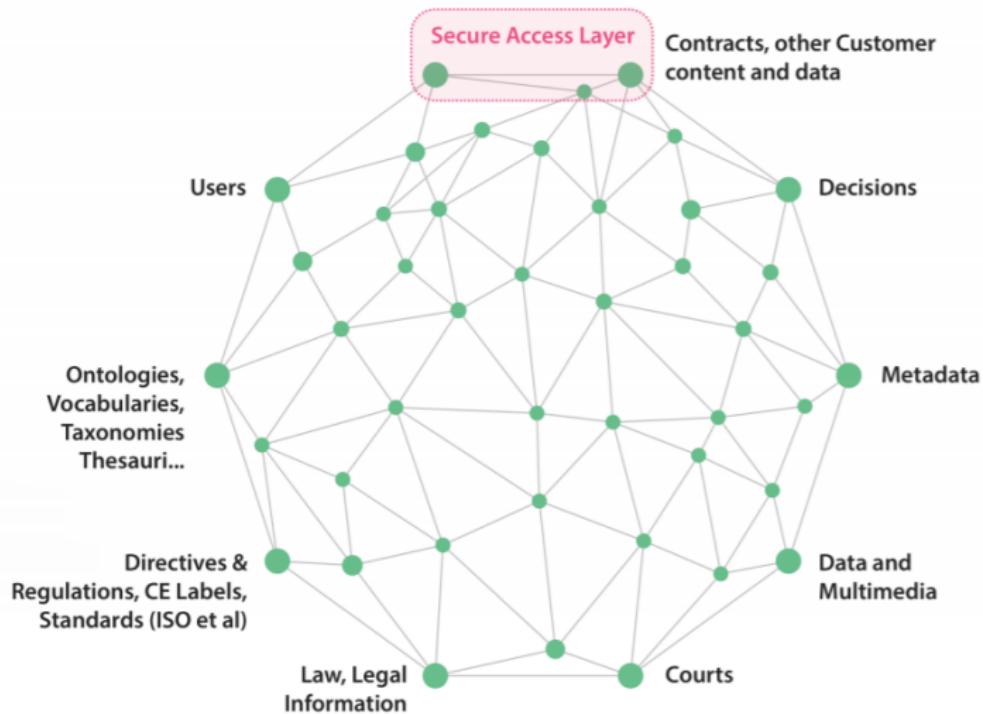
Fact Inference¹²

- ▶ Further data enrichment and data completion
 - ▶ Internal: Dominant type and label
 - ▶ External: Search engine method for enriching social links.

¹²Y. Gao, et al. "Building a large-scale, accurate and fresh knowledge graph." KDD-2018, Tutorial 39 (2018).

Towards a Legal Knowledge Graph

Our Vision of Legal KG¹³



¹³V. Mireles and M. Kaltenböck. "Building the Legal Knowledge Graph". W3C Workshop, 2018

Towards a Legal Knowledge Graph

Why Do We Build a Legal KG

- ▶ Cross-jurisdictional search

Towards a Legal Knowledge Graph

Why Do We Build a Legal KG

- ▶ Cross-jurisdictional search
- ▶ Support the comparative analyses of court decisions and different legal interpretations of legislation

Towards a Legal Knowledge Graph

Why Do We Build a Legal KG

- ▶ Cross-jurisdictional search
- ▶ Support the comparative analyses of court decisions and different legal interpretations of legislation
- ▶ Enables the evolution of legislation and jurisdiction

Towards a Legal Knowledge Graph

Why Do We Build a Legal KG

- ▶ Cross-jurisdictional search
- ▶ Support the comparative analyses of court decisions and different legal interpretations of legislation
- ▶ Enables the evolution of legislation and jurisdiction
- ▶ Interlink legal knowledge with external knowledge bases

Towards a Legal Knowledge Graph

Why Do We Build a Legal KG

- ▶ Cross-jurisdictional search
- ▶ Support the comparative analyses of court decisions and different legal interpretations of legislation
- ▶ Enables the evolution of legislation and jurisdiction
- ▶ Interlink legal knowledge with external knowledge bases
- ▶ Legal argument mining, reasoning, and summarization

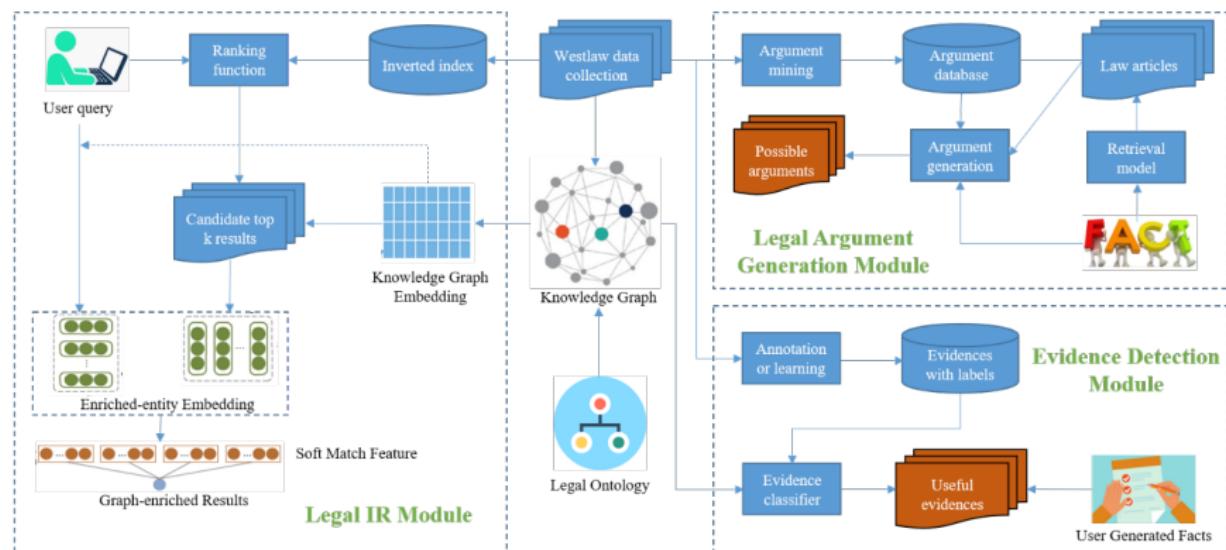
Towards a Legal Knowledge Graph

Why Do We Build a Legal KG

- ▶ Cross-jurisdictional search
- ▶ Support the comparative analyses of court decisions and different legal interpretations of legislation
- ▶ Enables the evolution of legislation and jurisdiction
- ▶ Interlink legal knowledge with external knowledge bases
- ▶ Legal argument mining, reasoning, and summarization
- ▶ Question answering

Towards a Legal Knowledge Graph

Build an intelligent legal system, for automated the information retrieval, evidence detection , and argument generation in legal applications.



Towards a Legal Knowledge Graph

Challenges

- ▶ The style and structure of legal text are complex, and high level domain knowledge are required ¹⁴.

¹⁴ B. Guido, L. Caro, and L. Humphreys. "Using classification to support legal knowledge engineers in the eunomos legal document management system." Fifth international workshop on Juris-informatics (JURISIN). 2011.

¹⁵ S. Pudaruth et al. "An innovative multi-segment strategy for the classification of legal judgments using the k-nearest neighbour classifier." Complex Intelligent Systems 4.1 (2018): 1-10.

Towards a Legal Knowledge Graph

Challenges

- ▶ The style and structure of legal text are complex, and high level domain knowledge are required ¹⁴.
- ▶ Legal text are often long, usually thousands of words. It may contain issues related to cross areas of laws, which makes information extraction difficult ¹⁴.

¹⁴ B. Guido, L. Caro, and L. Humphreys. "Using classification to support legal knowledge engineers in the eunomos legal document management system." Fifth international workshop on Juris-informatics (JURISIN). 2011.

¹⁵ S. Pudaruth et al. "An innovative multi-segment strategy for the classification of legal judgments using the k-nearest neighbour classifier." Complex Intelligent Systems 4.1 (2018): 1-10.

Towards a Legal Knowledge Graph

Challenges

- ▶ The style and structure of legal text are complex, and high level domain knowledge are required ¹⁴.
- ▶ Legal text are often long, usually thousands of words. It may contain issues related to cross areas of laws, which makes information extraction difficult ¹⁴.
- ▶ legal cases in different categories may use the common descriptions of an event, which makes it difficult for semantic understanding ¹⁵.

¹⁴ B. Guido, L. Caro, and L. Humphreys. "Using classification to support legal knowledge engineers in the eunomos legal document management system." Fifth international workshop on Juris-informatics (JURISIN). 2011.

¹⁵ S. Pudaruth et al. "An innovative multi-segment strategy for the classification of legal judgments using the k-nearest neighbour classifier." Complex Intelligent Systems 4.1 (2018): 1-10.

Towards a Legal Knowledge Graph

Challenges

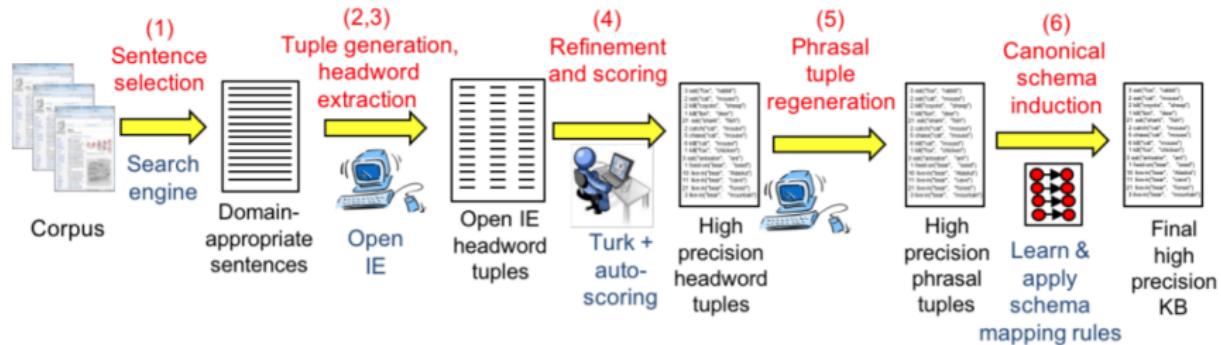
- ▶ The style and structure of legal text are complex, and high level domain knowledge are required ¹⁴.
- ▶ Legal text are often long, usually thousands of words. It may contain issues related to cross areas of laws, which makes information extraction difficult ¹⁴.
- ▶ legal cases in different categories may use the common descriptions of an event, which makes it difficult for semantic understanding ¹⁵.
- ▶ The lacking of automated tools in legal field.

¹⁴ B. Guido, L. Caro, and L. Humphreys. "Using classification to support legal knowledge engineers in the eunomos legal document management system." Fifth international workshop on Juris-informatics (JURISIN). 2011.

¹⁵ S. Pudaruth et al. "An innovative multi-segment strategy for the classification of legal judgments using the k-nearest neighbour classifier." Complex Intelligent Systems 4.1 (2018): 1-10.

Towards a Legal Knowledge Graph

How to Build a Legal KG - Our Solution¹⁶



¹⁶ M. Dalvi, N. Tandon, and P. Clark. "Domain-targeted, high precision knowledge extraction." *Transactions of the Association for Computational Linguistics* 5 (2017): 233-246.

Towards a Legal Knowledge Graph

How to Build a Legal KG - Our Solution

- ▶ Corpus: 650 millions of US Case law documents.

¹⁷ <https://lexpredict-lexnlp.readthedocs.io/en/latest/>

¹⁸ <https://nlp.stanford.edu/software/openie.html>

Towards a Legal Knowledge Graph

How to Build a Legal KG - Our Solution

- ▶ Corpus: 650 millions of US Case law documents.
- ▶ Sentence Selection: LexNLP¹⁷ was used for sentence splitting.

¹⁷ <https://lexpredict-lexnlp.readthedocs.io/en/latest/>

¹⁸ <https://nlp.stanford.edu/software/openie.html>

Towards a Legal Knowledge Graph

How to Build a Legal KG - Our Solution

- ▶ Corpus: 650 millions of US Case law documents.
- ▶ Sentence Selection: LexNLP ¹⁷ was used for sentence splitting.
- ▶ Tuple Generation: OpenIE ¹⁸ was used to extract an initial set of triples from the sentences.

¹⁷ <https://lexpredict-lexnlp.readthedocs.io/en/latest/>

¹⁸ <https://nlp.stanford.edu/software/openie.html>

Towards a Legal Knowledge Graph

How to Build a Legal KG - Our Solution

- ▶ Corpus: 650 millions of US Case law documents.
- ▶ Sentence Selection: LexNLP ¹⁷ was used for sentence splitting.
- ▶ Tuple Generation: OpenIE ¹⁸ was used to extract an initial set of triples from the sentences.
- ▶ Refinement and Scoring: Feature-based machine learning method based on partial tuples scored manually.

¹⁷ <https://lexpredict-lexnlp.readthedocs.io/en/latest/>

¹⁸ <https://nlp.stanford.edu/software/openie.html>

Towards a Legal Knowledge Graph

How to Build a Legal KG - Our Solution

- ▶ Corpus: 650 millions of US Case law documents.
- ▶ Sentence Selection: LexNLP ¹⁷ was used for sentence splitting.
- ▶ Tuple Generation: OpenIE ¹⁸ was used to extract an initial set of triples from the sentences.
- ▶ Refinement and Scoring: Feature-based machine learning method based on partial tuples scored manually.
- ▶ Relation Canonicalization: Identify the equivalent or similar relations, and map them to a canonical generalized relation.

¹⁷ <https://lexpredict-lexnlp.readthedocs.io/en/latest/>

¹⁸ <https://nlp.stanford.edu/software/openie.html>

Towards a Legal Knowledge Graph

How to Build a Legal KG - Our Solution

- ▶ Corpus: 650 millions of US Case law documents.
- ▶ Sentence Selection: LexNLP ¹⁷ was used for sentence splitting.
- ▶ Tuple Generation: OpenIE ¹⁸ was used to extract an initial set of triples from the sentences.
- ▶ Refinement and Scoring: Feature-based machine learning method based on partial tuples scored manually.
- ▶ Relation Canonicalization: Identify the equivalent or similar relations, and map them to a canonical generalized relation.
- ▶ Triple quality evaluation

¹⁷ <https://lexpredict-lexnlp.readthedocs.io/en/latest/>

¹⁸ <https://nlp.stanford.edu/software/openie.html>

Knowledge Graph Quality Evaluation

Why does KG quality matter? ¹⁹

- ▶ "Garbage in, garbage out"

¹⁹ H. Chen, G. Cao, J. Chen, J. Ding, "A Practical Framework for Evaluating the Quality of Knowledge Graph", CCKS 2019

Knowledge Graph Quality Evaluation

Why does KG quality matter? ¹⁹

- ▶ "Garbage in, garbage out"
- ▶ Foundation of KG-based Applications

¹⁹ H. Chen, G. Cao, J. Chen, J. Ding, "A Practical Framework for Evaluating the Quality of Knowledge Graph", CCKS 2019

Knowledge Graph Quality Evaluation

Why does KG quality matter? ¹⁹

- ▶ "Garbage in, garbage out"
- ▶ Foundation of KG-based Applications
- ▶ Data quality and Information Quality

¹⁹ H. Chen, G. Cao, J. Chen, J. Ding, "A Practical Framework for Evaluating the Quality of Knowledge Graph", CCKS 2019

Knowledge Graph Quality Evaluation

Why does KG quality matter? ¹⁹

- ▶ "Garbage in, garbage out"
- ▶ Foundation of KG-based Applications
- ▶ Data quality and Information Quality
- ▶ Intrinsic: Syntactic validity, semantic accuracy, consistence, correct, completeness

¹⁹ H. Chen, G. Cao, J. Chen, J. Ding, "A Practical Framework for Evaluating the Quality of Knowledge Graph", CCKS 2019

Knowledge Graph Quality Evaluation

Why does KG quality matter? ¹⁹

- ▶ "Garbage in, garbage out"
- ▶ Foundation of KG-based Applications
- ▶ Data quality and Information Quality
- ▶ Intrinsic: Syntactic validity, semantic accuracy, consistence, correct, completeness
- ▶ "Fit for purpose": Quality evaluation should be designed based on applications

¹⁹ H. Chen, G. Cao, J. Chen, J. Ding, "A Practical Framework for Evaluating the Quality of Knowledge Graph", CCKS 2019

Knowledge Graph Quality Evaluation

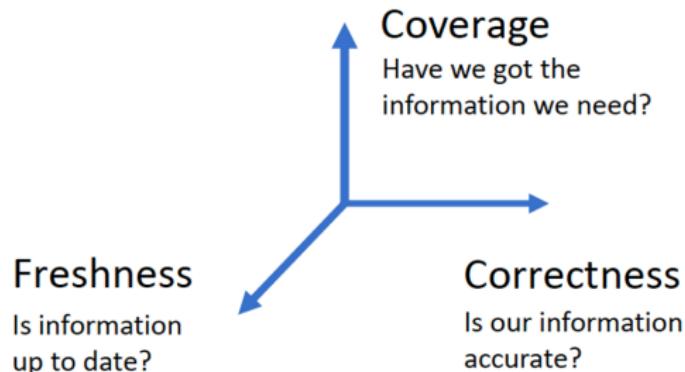
Why does KG quality matter? ¹⁹

- ▶ "Garbage in, garbage out"
- ▶ Foundation of KG-based Applications
- ▶ Data quality and Information Quality
- ▶ Intrinsic: Syntactic validity, semantic accuracy, consistence, correct, completeness
- ▶ "Fit for purpose": Quality evaluation should be designed based on applications
- ▶ A **practical** knowledge graph quality evaluation framework is therefore needed!

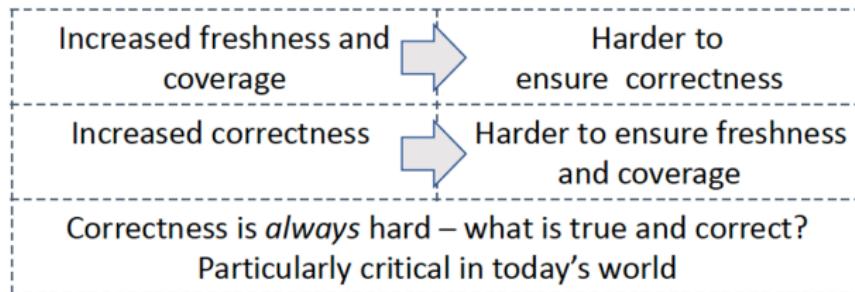
¹⁹ H. Chen, G. Cao, J. Chen, J. Ding, "A Practical Framework for Evaluating the Quality of Knowledge Graph", CCKS 2019

Knowledge Graph Quality Evaluation

Basic Quality Requirements for a KG



Three forces in constant conflict:



Knowledge Graph Quality Evaluation

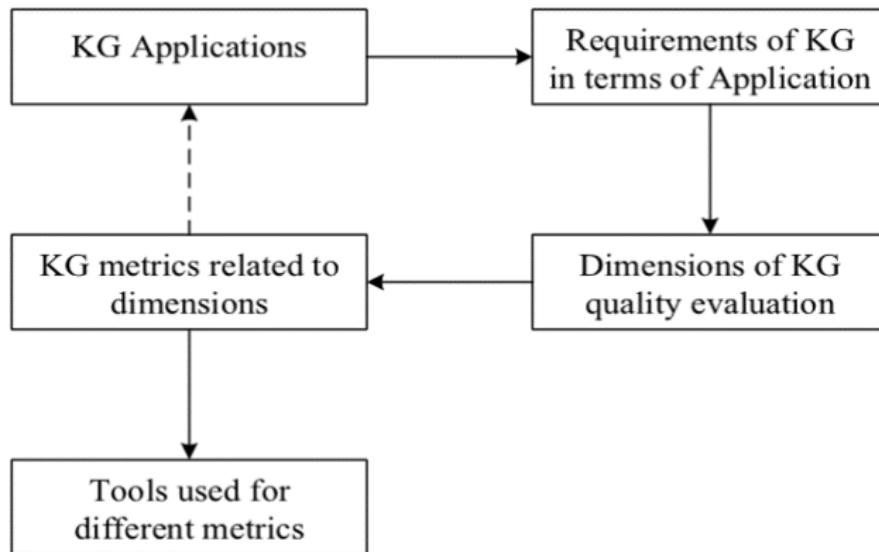
Comprehensive KG Quality Requirements²⁰

| | | | |
|-------------------------|---------------------------|------------------------------|----------------------------|
| 1. accessibility | 2. accuracy (veracity) | 3. appropriate amount | 4. believability |
| 5. completeness | 6. concise representation | 7. consistent representation | 8. cost-effectiveness |
| 9. easy of manipulating | 10. easy of operation | 11. easy of understanding | 12. flexibility |
| 13. free-of-error | 14. interoperability | 15. objectivity | 16. relevancy |
| 17. reputation | 18. security | 19. timeliness (velocity) | 20. traceability |
| 21. understandability | 22. value-added | 23. variety | **fitness for use** |

²⁰ Elias and Umutcan. "Building a large-scale, accurate and fresh knowledge graph." Semantics Conference 2019, Tutorial.

Knowledge Graph Quality Evaluation

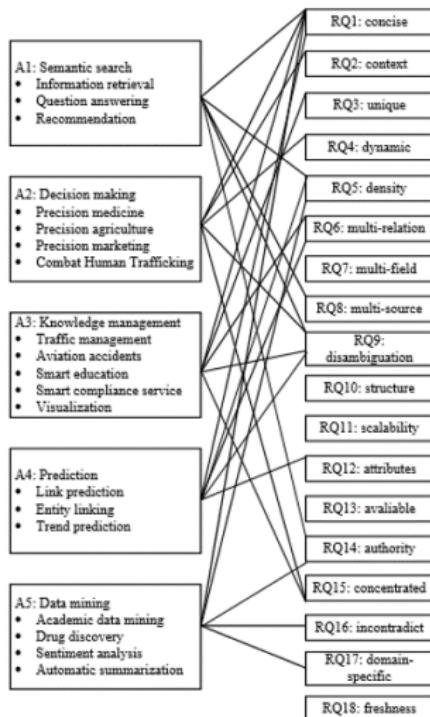
Knowledge graph quality evaluation framework ²¹



²¹ H. Chen, G. Cao, J. Chen, J. Ding, "A Practical Framework for Evaluating the Quality of Knowledge Graph", CCKS 2019

Knowledge Graph Quality Evaluation

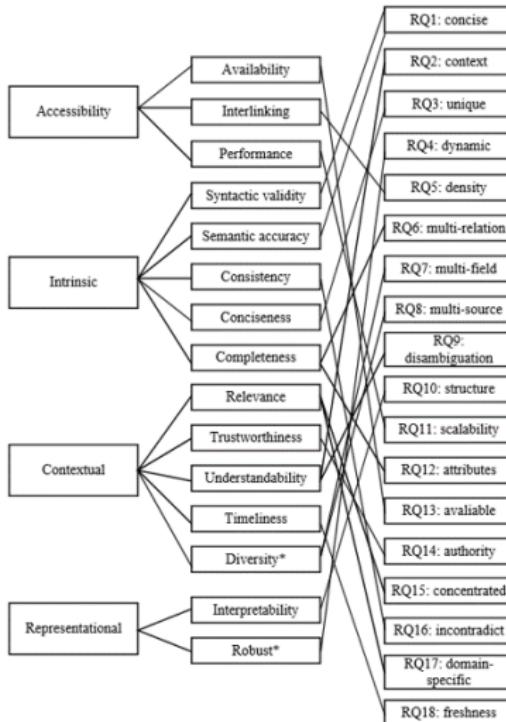
KG quality requirements mapped to applications²²



²²H. Chen, G. Cao, J. Chen, J. Ding, "A Practical Framework for Evaluating the Quality of Knowledge Graph", CCKS 2019

Knowledge Graph Quality Evaluation

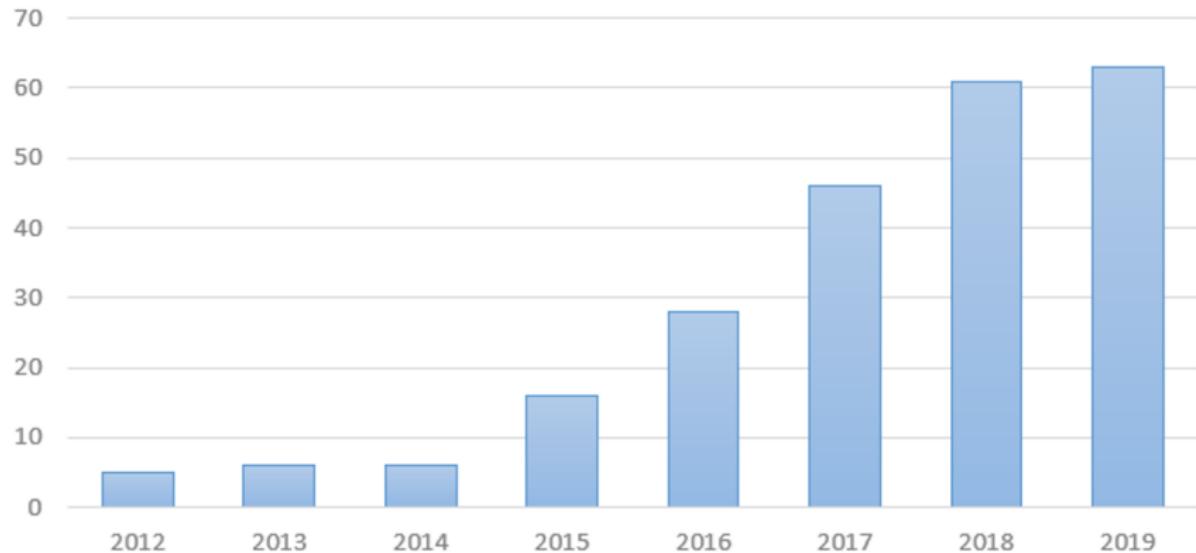
KG quality requirements mapped to quality dimensions²³



²³H. Chen, G. Cao, J. Chen, J. Ding, "A Practical Framework for Evaluating the Quality of Knowledge Graph", CCKS 2019

KG Applications

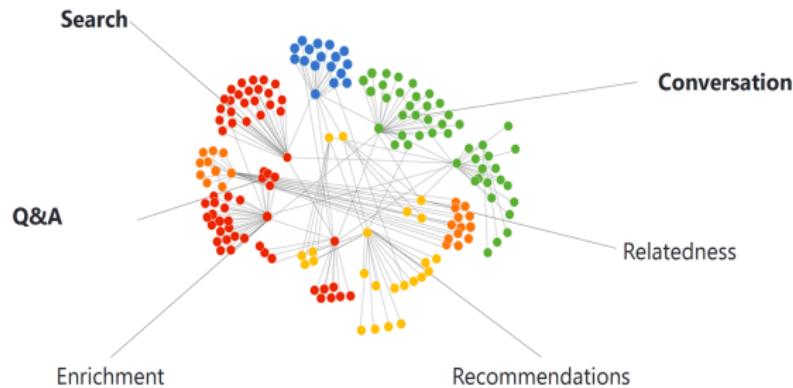
Significant Publications on Applications of KGs



KG Applications

Use Cases of Knowledge Graphs ²⁴

- ▶ Information Retrieval, Search
- ▶ Recommendation
- ▶ Questions Answering
- ▶ Medical Text Analysis
- ▶ Conversation
- ▶ ...



²⁴Y. Gao, et al. "Building a large-scale, accurate and fresh knowledge graph." KDD-2018, Tutorial 39 (2018).

KG Applications

Use Case: Search²⁵

- Leads to answers instead of web pages
 - Knowledge card provides rich background
 - Identify the entity from the search key word
 - Give answers directly

Google search results for "how old is trump". The search bar shows the query. Below it are filters: All, Images, News, Videos, Maps, More, Settings, Tools. A snippet says "About 251,000,000 results (0.54 seconds)". The main result is a knowledge card for Donald Trump:

Donald Trump / Age

70 years
June 14, 1946

Melania Trump Spouse 46 years
Hillary Clinton 69 years
Barack Obama 55 years

A small image of Donald Trump is shown next to the age information.

Google search results for "trump". The search bar shows the query. Below it are filters: All, Images, News, Videos, Maps, More, Settings, Tools. A snippet says "About 1,040,000,000 results (0.54 seconds)".

Donald Trump - Wikipedia
https://en.wikipedia.org/w/index.php?title=Trump&oldid=900000000
Donald John Trump is an American businessman, television personality, politician, and the 45th President of the United States. Born and...

Donald J. Trump for President: Make America Great Again!
https://www.donaldjtrump.com/
For 30+ years, Donald J. Trump is the very definition of the American success story, continually setting the standard for excellence in business, real estate and...

President Donald Trump: News, Twitter Updates and Quotes | Daily ...
https://www.dailymail.co.uk/trump/
Get the latest news on U.S. President Donald Trump's inauguration, administration, Twitter and investigate policy you need on the MailOnline homepage.

A Politicologist Analyzes Donald Trump's Personality - "The Atlantic"
https://www.theatlantic.com/politics/archive/2017/01/a-politician-analyzes-donald-trumps-personality/489777/
Politicians. Biographers. Journalists—psychologists investigate how Trump's authoritarian personality might drive his political policies.

Donald Trump | US News | The Gazette
https://www.gazette.com/trump/
Covering news, analysis, and commentary about the first truly business-run and political...

Donald J. Trump | Facebook
https://www.facebook.com/donaldjtrump/
Donald J. Trump, New York, NY 232,000 likes · 209,012 talking about this. This is the official Facebook page for Donald J. Trump.

Donald J. Trump | Facebook
https://www.facebook.com/donaldjtrump/
Donald J. Trump, New York, NY 232,000 likes · 209,012 talking about this. This is the official Facebook page for Donald J. Trump.

Trump News – Trump News
https://trumptv.psu.edu/
Throughout his first weeks in office, President Donald J. Trump has continued to advance the agenda he set out to win, while in private, working for...

Searches related to Trump

| | |
|-----------------------------|------------------------|
| trump poll | national trump |
| trump trump | donald trump education |
| donald trump on guns | donald trump biography |
| donald trump on immigration | donald trump facts |

People who also search for

| | | | | | | | | | |
|--------------|--------------|---------------|-----------------|----------------|------------------|------------|----------------|-----------------|--------------|
| Donald Trump | Barack Obama | Melania Trump | Chelsea Clinton | Michelle Obama | The First Family | First Lady | First Daughter | First Gentleman | First Family |
|--------------|--------------|---------------|-----------------|----------------|------------------|------------|----------------|-----------------|--------------|

25

Y. Xiao. "Introduction to Knowledge Graphs". Fudan University, 2017.

KG Applications

Use Case: Search

- ▶ Offer related knowledge that is related to searched results.

Watch now actions for movie entities

Finding Dory (PG - 2016 - 1hr 43min - Animation/Adventure) 7.4/10 94% 77 Disney Pixar's "Finding Dory" welcome back to the big screen everyone's favorite forgetful blue tang Dory, who's living happily in the reef next to Marlin and Nemo. When Dory suddenly remembers that she has a family out there who may be looking for her, the fish takes off on a life-changing adventure across the ocean to California's prestigious Marine Life Inst. +

WATCH NOW

Amazon Watch iTunes Watch Netflix Watch

Recently viewed:

Coco Star Wars: Episode V - The Empire... Return of the Jedi Star Wars: The Force Awakens Star Wars: The Last Jedi Star Wars: The Force Awakens

See all (6+)

TV listings for TV show entities

The Big Bang Theory

CBS - Tue 1:30/12:30 AM CT 8.3/10 ★★★★☆ 8.9/10 ★★★★★ A woman who moves into an apartment across the hall from two brilliant but socially awkward physicists shows them how little they know about life outside of the laboratory.

Official website Wikipedia YouTube Facebook Twitter

First episode: Sep 24, 2007
Episode duration: 22 minutes
Creators: Chuck Lorre · Bill Prady
Theme song: Big Bang Theory Theme
Origin: United States
Producers: Chuck Lorre · Bill Prady · Steven Molaro · Faye Olshansky · Belyeu · Lee Aronsohn · Eddie Gorodetsky · Robert Cohan +

On TV

All times are Eastern Time

| Thu, 11/30, 8 PM | CBS | S11, Ep09 | NEW |
|------------------|-----|-----------|-----|
| Thu, 12/7, 8 PM | CBS | S11, Ep10 | NEW |
| Today, 7:30 PM | TBS | S3, Ep10 | |
| Today, 8 PM | TBS | S3, Ep11 | |
| Today, 8:30 PM | TBS | S3, Ep12 | |

See more ▾

Job insights for companies

Boeing

The Boeing Company is an American multinational corporation that designs, manufactures, and sells aerospace, defense, security, and space products and services. The company also provides leasing and product support services. Boeing is among the largest gl... +

Official website Wikipedia Twitter Facebook LinkedIn

Stock price: BA (NYSE) \$28.22 ▼-1.72 (-6.02%)
Div 05, 12-14 PM EST 30 min delayed
Customer service: +1-208-951-1131
Founded: Jul 15, 1916 Seattle, WA
Revenue: \$64.57 billion USD (2018)
Headquarters: Chicago, IL
CEO: Dennis Muilenburg (Since 2015)

Related people

William Boeing Founder Dennis Muilenburg CEO James Mohrley Former CEO Raymond Conner Harry Stonecipher

Job insights

| Title | Openings | Est. base salary |
|----------------------|----------|------------------|
| Software Engineer | 227 | \$95-120k |
| System Engineer | 180 | \$131-149k |
| Aircraft Mechanic | 80 | \$59-73k |
| All other job titles | 257 | |

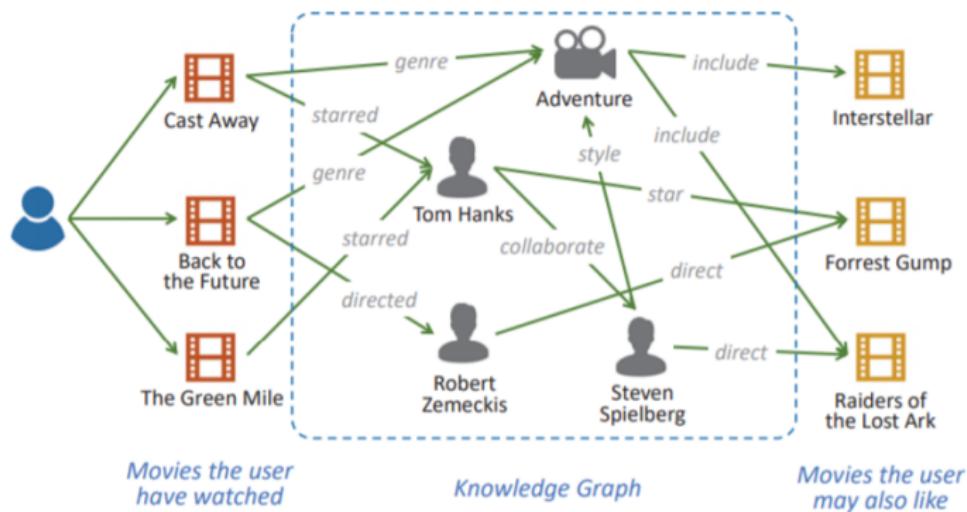
People also search for

Airbus Lockheed Martin Northrop Grumman Raytheon Airbus Group

KG Applications

Use Case: Search

- ▶ Recommend semantically related content.
 - ▶ Example: What a search engine should recommend if a user search for “flu shot”?
- ▶ Recommend conceptually consistent content.
 - ▶ Example: What Netflix should recommend if a user search for “Star War”?



KG Applications

Use Case: Question Answering

Fact Answering

how many students attend UNT denton tx

All Images Videos Maps News Shopping | My saves

107,000 Results Any time ▾

UNT is the choice of more than 39,000 **students** and is a student-centered public research university. As one of Texas' largest universities, UNT offers 105 bachelor's, 88 master's and 37 doctoral degree programs, many nationally and internationally recognized.

Parents and Family | University of North Texas
www.unt.edu/parents-family

Was this helpful? thumb up thumb down

About UNT | University of North Texas
<https://www.unt.edu/about-unt>

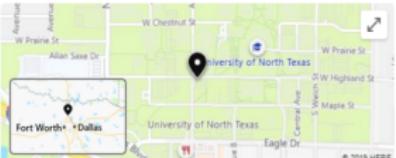
Programs Mission Academics Research Philosophy A >

As one of the nation's largest universities, we offer 105 bachelor's, 88 master's and 37 doctoral degree programs. By providing access, welcoming diversity and strengthening our collaborations with our many educational, business and community partners, as well as building new partnerships across the globe, UNT's faculty and staff work each day to prepare students for the challenges they will meet in our changing world.

See more on unt.edu

University of North Texas - Wikipedia
https://en.wikipedia.org/wiki/University_of_North_Texas

The University of North Texas (UNT) is a public research university in Denton, Texas. Eleven colleges, two schools, an early admissions math and science academy for exceptional high-school-age students from across the state, and a library system comprising the university's main



W Chestnut St
Allan Saxe Dr
W Prairie St
University of North Texas
Eagle Dr
S Union St
Central Ave
Maple St
W Highland St
W Prairie St
Airport Rd
Austin Rd
W Prairie St

© 2019 HERE

University of North Texas
Public University

The University of North Texas is a public research university in Denton, Texas. Eleven colleges, two sch...

UNT +

Directions Official site

Acceptance rate: 71.00% (2018-19)
Tuition: in-state \$11,251 USD / out-of-state \$23,911 USD (2019)
Undergraduates: 31,347 (2019)
Enrollment: 38,087 (2019)
Mascot: University of North Texas Scrappy
Founded: 1890

KG Applications

Use Case: Question Answering

Knowledge-based Q&A

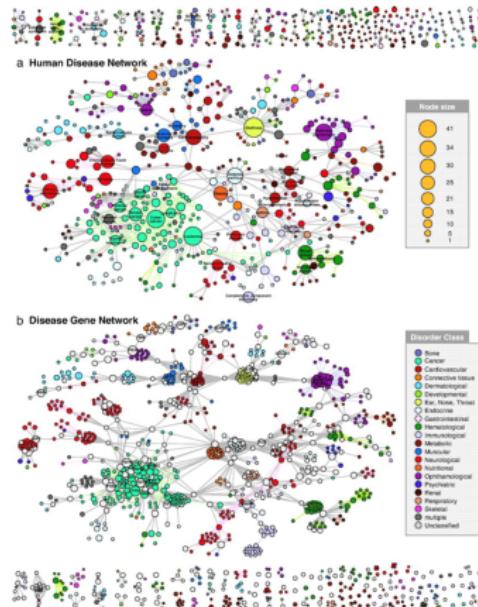
The screenshot shows a search results page from a web browser. The search query is "most touchdown passes". The results are filtered under the heading "Microsoft Show business results >". The top result is Peyton Manning, described as "Leading NFL for passing touchdowns" with 539. Below him is a bar chart comparing five NFL players:

| Player | Touchdown Passes |
|----------------|------------------|
| Brett Favre | 508 |
| Tom Brady | 428 |
| Drew Brees | 428 |
| Dan Marino | 420 |
| Fran Tarkenton | 342 |

A large, semi-transparent knowledge graph is overlaid on the right side of the page. It consists of numerous purple nodes connected by lines, forming a complex network structure. A grey arrow points from the bottom left towards the graph.

KG Applications

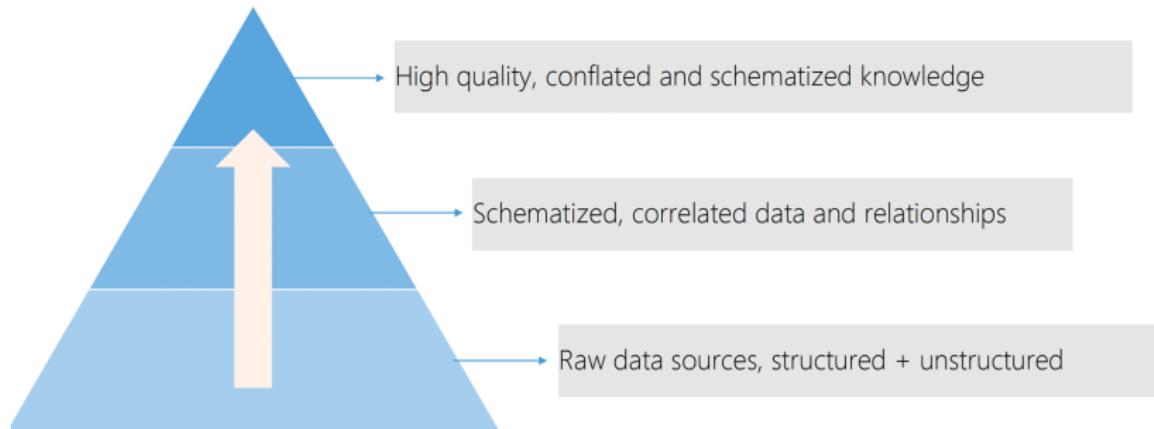
Use Case: Medical Text Analytics²⁶



²⁶Y. Xiao. "Introduction to Knowledge Graphs". Fudan University, 2017.

Summary

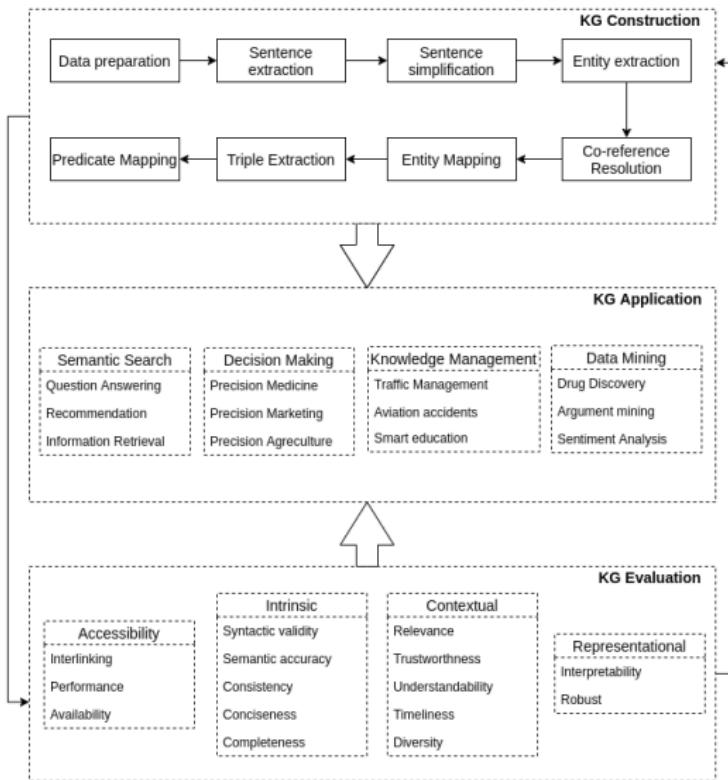
KG for AI: ML + NLP + Conflation + Inference ²⁷



²⁷ Y. Gao, et al. "Building a large-scale, accurate and fresh knowledge graph." KDD-2018, Tutorial 39 (2018).

Summary

Combine KG Construction, Evaluation, with Application



Summary

Future Directions of KG

- ▶ All KG creation methods have their advantages and disadvantages, how to set the most appropriate method based on downstream tasks deserved attentions²⁸.

²⁸P. Bonatti, et al. "Knowledge graphs: new directions for knowledge representation on the semantic web (Dagstuhl seminar 18371)." Schloss Dagstuhl-Leibniz-Zentrum fuer Informatik, 2019.

²⁹L. Yao, et al. "KG-BERT: BERT for Knowledge Graph Completion." arXiv preprint arXiv:1909.03193 (2019).

Summary

Future Directions of KG

- ▶ All KG creation methods have their advantages and disadvantages, how to set the most appropriate method based on downstream tasks deserved attentions²⁸.
- ▶ A unified framework integrating different reasoning paradigms needs to be formalized²⁸.

²⁸P. Bonatti, et al. "Knowledge graphs: new directions for knowledge representation on the semantic web (Dagstuhl seminar 18371)." Schloss Dagstuhl-Leibniz-Zentrum fuer Informatik, 2019.

²⁹L. Yao, et al. "KG-BERT: BERT for Knowledge Graph Completion." arXiv preprint arXiv:1909.03193 (2019).

Summary

Future Directions of KG

- ▶ All KG creation methods have their advantages and disadvantages, how to set the most appropriate method based on downstream tasks deserved attentions²⁸.
- ▶ A unified framework integrating different reasoning paradigms needs to be formalized²⁸.

²⁸P. Bonatti, et al. "Knowledge graphs: new directions for knowledge representation on the semantic web (Dagstuhl seminar 18371)." Schloss Dagstuhl-Leibniz-Zentrum fuer Informatik, 2019.

²⁹L. Yao, et al. "KG-BERT: BERT for Knowledge Graph Completion." arXiv preprint arXiv:1909.03193 (2019).

Summary

Future Directions of KG

- ▶ All KG creation methods have their advantages and disadvantages, how to set the most appropriate method based on downstream tasks deserved attentions²⁸.
- ▶ A unified framework integrating different reasoning paradigms needs to be formalized²⁸.
- ▶ Use language models like BERT in conjunction with knowledge graph embeddings is the future work to enhance semantic representation²⁹. GNN might be a good answer.

²⁸P. Bonatti, et al. "Knowledge graphs: new directions for knowledge representation on the semantic web (Dagstuhl seminar 18371)." Schloss Dagstuhl-Leibniz-Zentrum fuer Informatik, 2019.

²⁹L. Yao, et al. "KG-BERT: BERT for Knowledge Graph Completion." arXiv preprint arXiv:1909.03193 (2019).

Summary

Future Directions of KG

- ▶ All KG creation methods have their advantages and disadvantages, how to set the most appropriate method based on downstream tasks deserved attentions²⁸.
- ▶ A unified framework integrating different reasoning paradigms needs to be formalized²⁸.
- ▶ Use language models like BERT in conjunction with knowledge graph embeddings is the future work to enhance semantic representation²⁹. GNN might be a good answer.
- ▶ ...

²⁸P. Bonatti, et al. "Knowledge graphs: new directions for knowledge representation on the semantic web (Dagstuhl seminar 18371)." Schloss Dagstuhl-Leibniz-Zentrum fuer Informatik, 2019.

²⁹L. Yao, et al. "KG-BERT: BERT for Knowledge Graph Completion." arXiv preprint arXiv:1909.03193 (2019).

Questions, Comments and Suggestions?

Thank you!