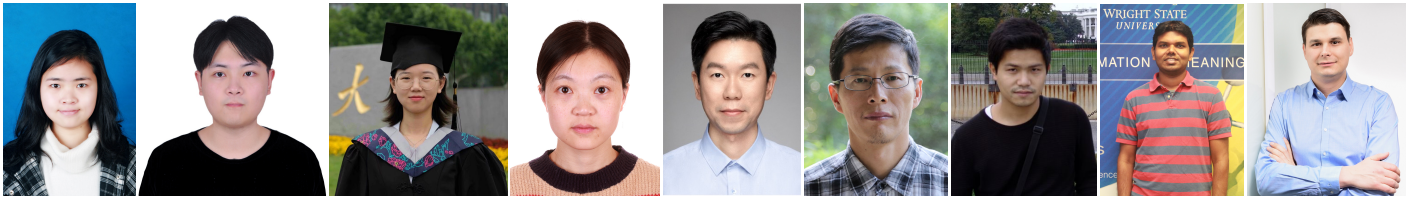
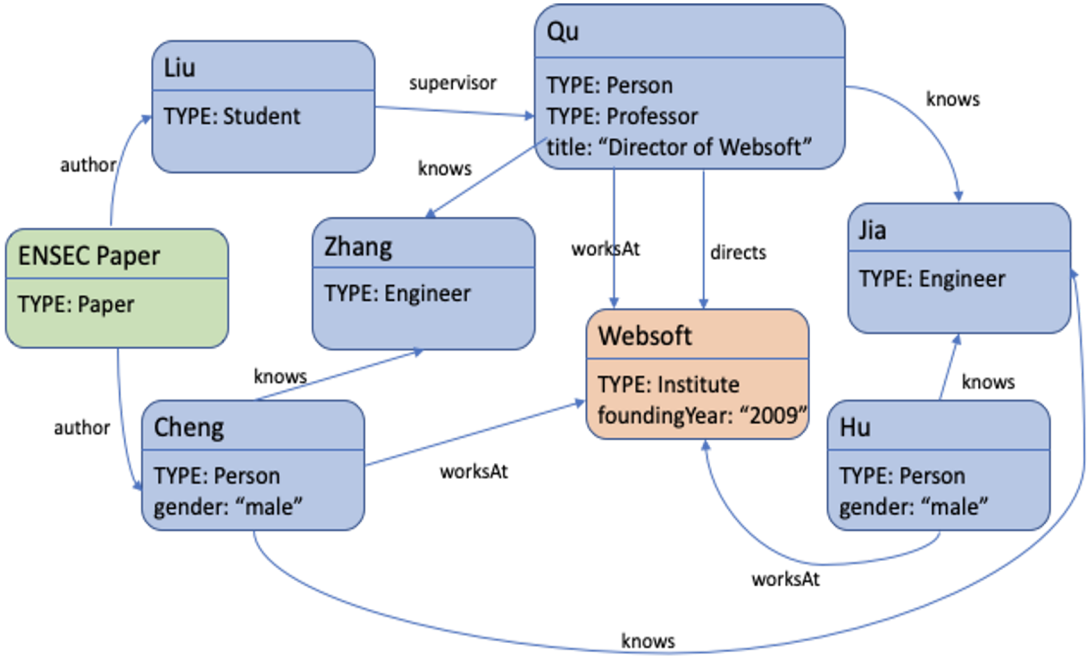
Entity Summarization: Where We Are and What Lies Ahead

Gong Cheng
Nanjing University

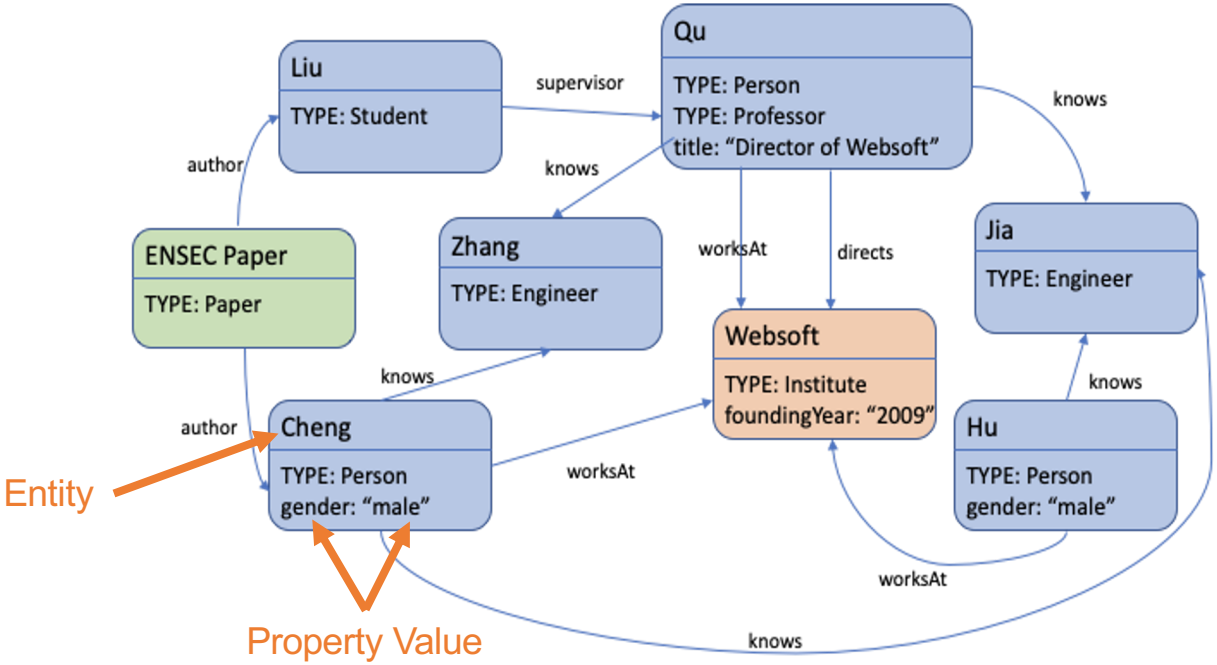
Presented at EEKE@JCDL, 09/30/2021



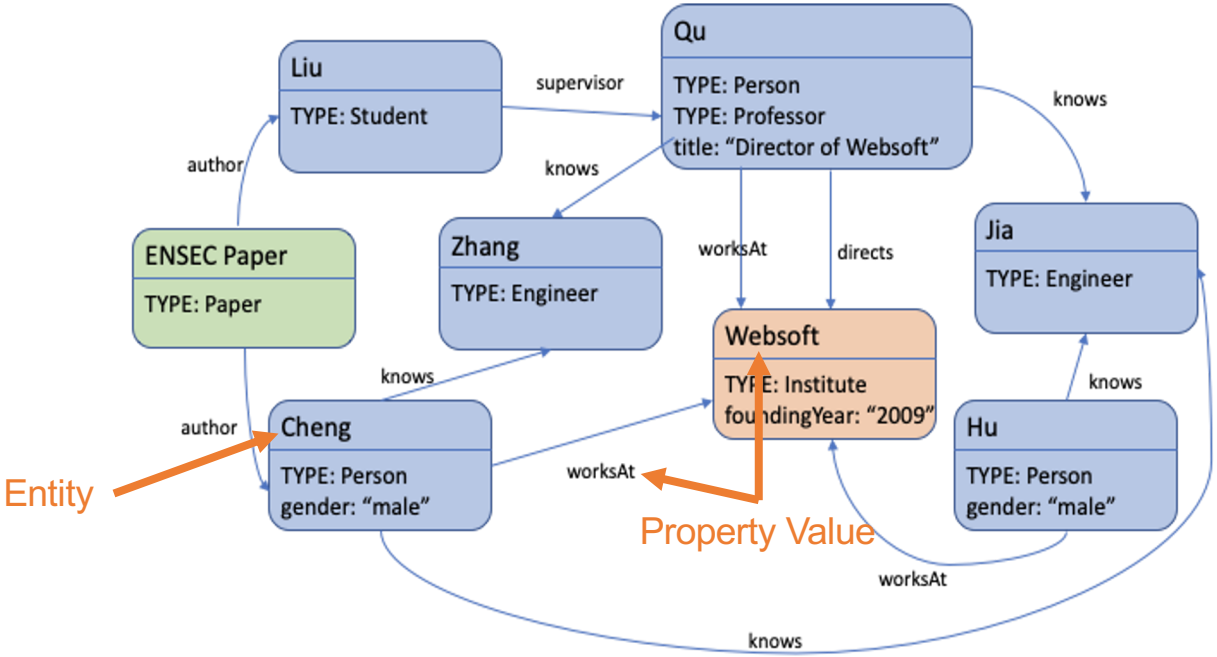
Knowledge Graph and Entities



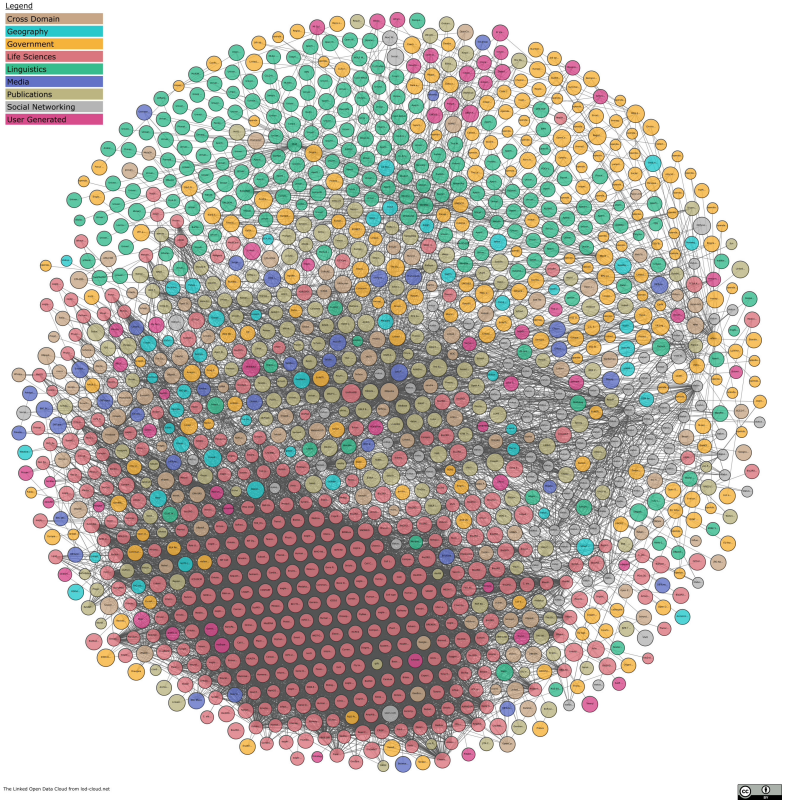
Knowledge Graph and Entities



Knowledge Graph and Entities

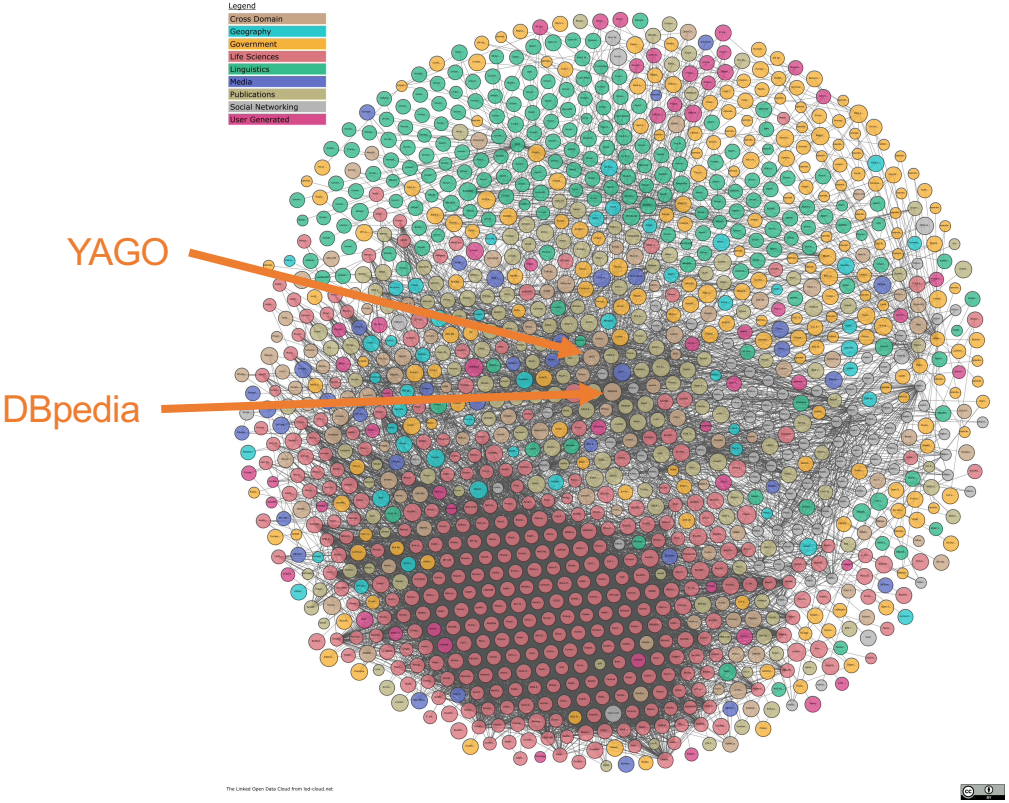


Knowledge Graphs on the Web



https://lod-cloud.net/

Knowledge Graphs on the Web



https://lod-cloud.net/

Illinois in DBpedia

```
@prefix dbo:    <http://dbpedia.org/ontology/> .
@prefix dbr:    <http://dbpedia.org/resource/> .
dbr:David_Ayer  dbo:birthPlace  dbr:Illinois .
dbr:Samuel_Hynes    dbo:wikiPageWikiLink    dbr:Illinois .
<http://dbpedia.org/resource/C._Denier_Warren>  dbo:wikiPageWikiLink    dbr:Illinois ;
    dbo:birthPlace  dbr:Illinois .
<http://dbpedia.org/resource/John_Humphrey_(Illinois_politician)>   dbo:deathPlace  dbr:Illinois .
@prefix dbp:    <http://dbpedia.org/property/> .
<http://dbpedia.org/resource/John_Humphrey_(Illinois_politician)>   dbp:deathPlace  dbr:Illinois .
dbr:John_Mahoney    dbo:deathPlace  dbr:Illinois .
<http://dbpedia.org/resource/Thomas_B._Dunstan> dbo:deathPlace  dbr:Illinois ;
    dbp:deathPlace  dbr:Illinois .
<http://dbpedia.org/resource/Barber\u2013Colman_Company>    dbo:wikiPageWikiLink    dbr:Illinois .
dbr:Basement_Revolver   dbo:wikiPageWikiLink    dbr:Illinois .
dbr:Beagle  dbo:wikiPageWikiLink    dbr:Illinois .
dbr:List_of_hospice_programs    dbo:wikiPageWikiLink    dbr:Illinois .
dbr:List_of_lesbian_feminist_organizations  dbo:wikiPageWikiLink    dbr:Illinois .
dbr:List_of_sports_teams_named_Spartans dbo:wikiPageWikiLink    dbr:Illinois .
<http://dbpedia.org/resource/Herbert_J._Tweedie>    dbo:wikiPageWikiLink    dbr:Illinois .
<http://dbpedia.org/resource/Lyon_&_Healy>  dbo:wikiPageWikiLink    dbr:Illinois .
```

...

```
    dbp:birthPlace  dbr:Illinois .
dbr:Mike_Foltynewicz    dbo:birthPlace  dbr:Illinois ;
    dbp:birthPlace  dbr:Illinois .
<http://dbpedia.org/resource/T._J._McFarland>   dbo:birthPlace  dbr:Illinois ;
    dbp:birthPlace  dbr:Illinois .
<http://dbpedia.org/resource/Fred_Cooke_(baseball)> dbo:birthPlace  dbr:Illinois .
<http://dbpedia.org/resource/Walter_Young_(American_football)>  dbp:birthPlace  dbr:Illinois .
dbr:Ed_Sedar    dbo:birthPlace  dbr:Illinois ;
    dbp:birthPlace  dbr:Illinois .
dbr:Danny_Jansen    dbo:birthPlace  dbr:Illinois ;
    dbp:birthPlace  dbr:Illinois .
```

Illinois in Wikipedia

Illinois

From Wikipedia, the free encyclopedia

This article is about the State of Illinois. For the river, see Illinois River. For other uses, see Illinois (disambiguation).

Illinois (/ˌɪlɪˈnɔɪ/ (listen) *IL-ə-NOY*) is a state in the Midwestern region of the United States. It has the fifth largest gross domestic product (GDP), the sixth largest population, and the 25th largest land area of all U.S. states. Illinois has been noted as a microcosm of the entire United States.[7] With Chicago in northeastern Illinois, small industrial cities and immense agricultural productivity in the north and center of the state, and natural resources such as coal, timber, and petroleum in the south, Illinois has a diverse economic base, and is a major transportation hub. The Port of Chicago connects the state to international ports via two main routes: from the Great Lakes, via the Saint Lawrence Seaway, to the Atlantic Ocean and from the Great Lakes to the Mississippi River, via the Illinois River, through the Illinois Waterway. The Mississippi River, the Ohio River, and the Wabash River form parts of the boundaries of Illinois. For decades, Chicago's O'Hare International Airport has been ranked as one of the world's busiest airports. Illinois has long had a reputation as a bellwether both in social and cultural terms[7] and, through the 1980s, in politics.

The capital of Illinois is Springfield, which is located in the central part of the state. Although today Illinois's largest population center is in its northeast, the state's European population grew first in the west as the French settled lands near the Mississippi River, when the region was known as Illinois Country and was part of New France. Following the American Revolutionary War, American settlers began arriving from Kentucky in the 1780s via the Ohio River, and the population grew from south to north. In 1818, Illinois achieved statehood. Following increased commercial activity in the Great Lakes after the construction of the Erie Canal, Chicago was incorporated in the 1830s on the banks of the Chicago River at one of the few natural harbors on the southern section of Lake Michigan.[8] John Deere's invention of the self-scouring steel plow turned Illinois's rich prairie into some of the world's most productive and valuable farmland, attracting immigrant farmers from Germany and Sweden. The Illinois and Michigan Canal (1848) made transportation between the Great Lakes and the Mississippi River valley faster and cheaper, and new railroads carried immigrants to new homes in the country's west and shipped commodity crops to the nation's east. The state became a transportation hub for the nation.[9]

By 1900, the growth of industrial jobs in the northern cities and coal mining in the central and southern areas attracted immigrants from Eastern and Southern Europe. Illinois was an important manufacturing center during both world wars. The Great Migration from the South established a large community of African Americans in the state, including Chicago, who founded the city's famous jazz and blues cultures.[10][11] Chicago, the center of the Chicago Metropolitan Area, is now recognized as a global city. Chicagoland, Chicago's metropolitan area, encompasses about 65% of the state's population. The most populous metropolitan areas outside the Chicago area include, Metro East (of Greater St. Louis), Peoria and Rockford.

Three U.S. presidents have been elected while living in Illinois: Abraham Lincoln, Ulysses S. Grant, and Barack Obama. Additionally, Ronald Reagan, whose political career was based in California, was born and raised in the state. Today, Illinois honors Lincoln with its official state slogan *Land of Lincoln*, which has been displayed on its license plates since 1954.[12][13] The state is the site of the Abraham Lincoln Presidential Library and Museum in Springfield and the future home of the Barack Obama Presidential Center in Chicago.

Contents [hide]

1 Etymology
2 History
  2.1 Geologic history
  2.2 Pre-European
  2.3 European exploration and settlement prior to 1800
  2.4 19th century
  2.5 20th century
  2.6 21st century
3 Geography
  3.1 Boundaries
  3.2 Topography
  3.3 Divisions
  3.4 Climate
4 Demographics
  4.1 2019 American Community Survey
  4.2 Birth data
  4.3 Urban areas
  4.4 Languages
  4.5 Religion
5 Economy

Illinois in Wikipedia

Illinois

From Wikipedia, the free encyclopedia

This article is about the State of Illinois. For the river, see Illinois River. For other uses, see Illinois (disambiguation).

Illinois (/ˌɪlɪˈnɔɪ/ (listen) *IL-ə-NOY*) is a state in the Midwestern region of the United States. It has the fifth largest gross domestic product (GDP), the sixth largest population, and the 25th largest land area of all U.S. states. Illinois has been noted as a microcosm of the entire United States.[7] With Chicago in northeastern Illinois, small industrial cities and immense agricultural productivity in the north and center of the state, and natural resources such as coal, timber, and petroleum in the south, Illinois has a diverse economic base, and is a major transportation hub. The Port of Chicago connects the state to international ports via two main routes: from the Great Lakes, via the Saint Lawrence Seaway, to the Atlantic Ocean and from the Great Lakes to the Mississippi River, via the Illinois River, through the Illinois Waterway. The Mississippi River, the Ohio River, and the Wabash River form parts of the boundaries of Illinois. For decades, Chicago's O'Hare International Airport has been ranked as one of the world's busiest airports. Illinois has long had a reputation as a bellwether both in social and cultural terms[7] and, through the 1980s, in politics.

The capital of Illinois is Springfield, which is located in the central part of the state. Although today Illinois's largest population center is in its northeast, the state's European population grew first in the west as the French settled lands near the Mississippi River, when the region was known as Illinois Country and was part of New France. Following the American Revolutionary War, American settlers began arriving from Kentucky in the 1780s via the Ohio River, and the population grew from south to north. In 1818, Illinois achieved statehood. Following increased commercial activity in the Great Lakes after the construction of the Erie Canal, Chicago was incorporated in the 1830s on the banks of the Chicago River at one of the few natural harbors on the southern section of Lake Michigan.[8] John Deere's invention of the self-scouring steel plow turned Illinois's rich prairie into some of the world's most productive and valuable farmland, attracting immigrant farmers from Germany and Sweden. The Illinois and Michigan Canal (1848) made transportation between the Great Lakes and the Mississippi River valley faster and cheaper, and new railroads carried immigrants to new homes in the country's west and shipped commodity crops to the nation's east. The state became a transportation hub for the nation.[9]

By 1900, the growth of industrial jobs in the northern cities and coal mining in the central and southern areas attracted immigrants from Eastern and Southern Europe. Illinois was an important manufacturing center during both world wars. The Great Migration from the South established a large community of African Americans in the state, including Chicago, who founded the city's famous jazz and blues cultures.[10][11] Chicago, the center of the Chicago Metropolitan Area, is now recognized as a global city. Chicagoland, Chicago's metropolitan area, encompasses about 65% of the state's population. The most populous metropolitan areas outside the Chicago area include, Metro East (of Greater St. Louis), Peoria and Rockford.

Three U.S. presidents have been elected while living in Illinois: Abraham Lincoln, Ulysses S. Grant, and Barack Obama. Additionally, Ronald Reagan, whose political career was based in California, was born and raised in the state. Today, Illinois honors Lincoln with its official state slogan *Land of Lincoln*, which has been displayed on its license plates since 1954.[12][13] The state is the site of the Abraham Lincoln Presidential Library and Museum in Springfield and the future home of the Barack Obama Presidential Center in Chicago.

Article Summary

Contents [hide]

1 Etymology
2 History
  2.1 Geologic history
  2.2 Pre-European
  2.3 European exploration and settlement prior to 1800
  2.4 19th century
  2.5 20th century
  2.6 21st century
3 Geography
  3.1 Boundaries
  3.2 Topography
  3.3 Divisions
  3.4 Climate
4 Demographics
  4.1 2019 American Community Survey
  4.2 Birth data
  4.3 Urban areas
  4.4 Languages
  4.5 Religion
5 Economy

Illinois in DBpedia

```
@prefix dbo:    <http://dbpedia.org/ontology/> .
@prefix dbr:    <http://dbpedia.org/resource/> .
dbr:David_Ayer  dbo:birthPlace  dbr:Illinois .
dbr:Samuel_Hynes    dbo:wikiPageWikiLink    dbr:Illinois .
<http://dbpedia.org/resource/C._Denier_Warren>  dbo:wikiPageWikiLink    dbr:Illinois ;
    dbo:birthPlace  dbr:Illinois .
<http://dbpedia.org/resource/John_Humphrey_(Illinois_politician)>   dbo:deathPlace  dbr:Illinois .
@prefix dbp:    <http://dbpedia.org/property/> .
<http://dbpedia.org/resource/John_Humphrey_(Illinois_politician)>   dbp:deathPlace  dbr:Illinois .
dbr:John_Mahoney    dbo:deathPlace  dbr:Illinois .
<http://dbpedia.org/resource/Thomas_B._Dunstan> dbo:deathPlace  dbr:Illinois ;
    dbp:deathPlace  dbr:Illinois .
<http://dbpedia.org/resource/Barber\u2013Colman_Company>    dbo:wikiPageWikiLink    dbr:Illinois .
dbr:Basement_Revolver   dbo:wikiPageWikiLink    dbr:Illinois .
dbr:Beagle  dbo:wikiPageWikiLink    dbr:Illinois .
dbr:List_of_hospice_programs    dbo:wikiPageWikiLink    dbr:Illinois .
dbr:List_of_lesbian_feminist_organizations  dbo:wikiPageWikiLink    dbr:Illinois .
dbr:List_of_sports_teams_named_Spartans dbo:wikiPageWikiLink    dbr:Illinois .
<http://dbpedia.org/resource/Herbert_J._Tweedie>    dbo:wikiPageWikiLink    dbr:Illinois .
<http://dbpedia.org/resource/Lyon_&_Healy>  dbo:wikiPageWikiLink    dbr:Illinois .
```

How can we summarize an entity description in KG?

```
    dbp:birthPlace  dbr:Illinois .
dbr:Mike_Foltynewicz    dbo:birthPlace  dbr:Illinois ;
    dbp:birthPlace  dbr:Illinois .
<http://dbpedia.org/resource/T._J._McFarland>   dbo:birthPlace  dbr:Illinois ;
    dbp:birthPlace  dbr:Illinois .
<http://dbpedia.org/resource/Fred_Cooke_(baseball)> dbo:birthPlace  dbr:Illinois .
<http://dbpedia.org/resource/Walter_Young_(American_football)>  dbp:birthPlace  dbr:Illinois .
dbr:Ed_Sedar    dbo:birthPlace  dbr:Illinois ;
    dbp:birthPlace  dbr:Illinois .
dbr:Danny_Jansen    dbo:birthPlace  dbr:Illinois ;
    dbp:birthPlace  dbr:Illinois .
```

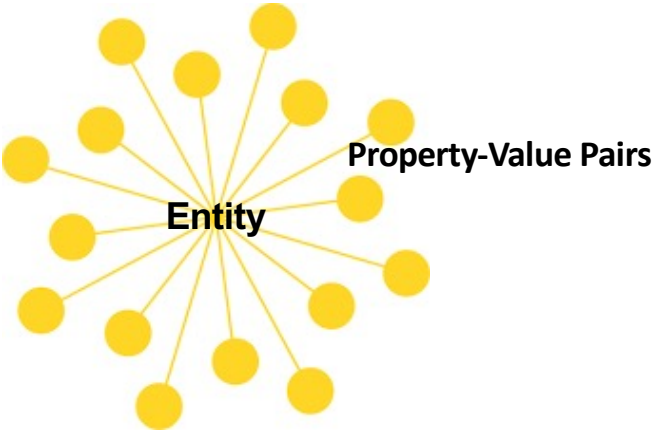
Outline

- Where We Are
  - What is entity summarization?
  - Technical features for entity summarization
  - Frameworks for feature combination
  - Deep learning based entity summarization
- What Lies Ahead

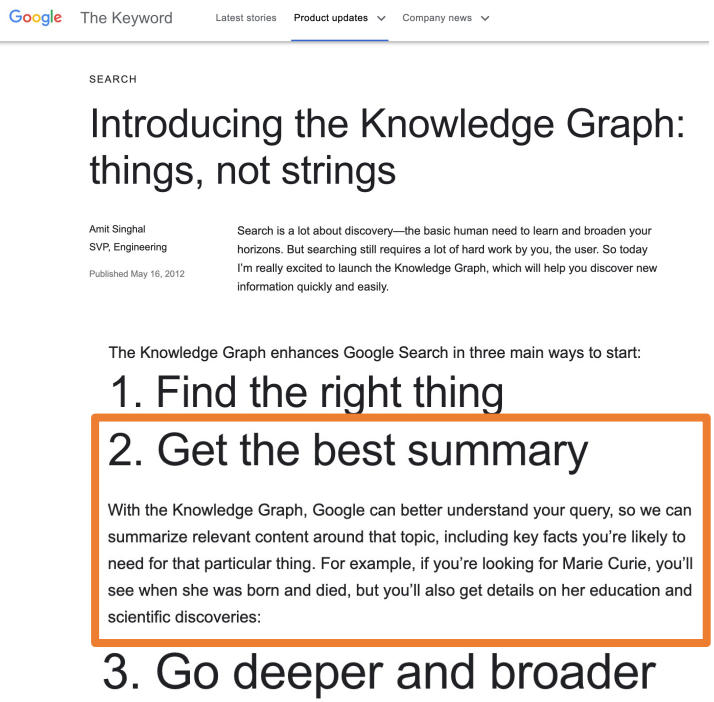
Outline

- Where We Are
  - What is entity summarization?
  - Technical features for entity summarization
  - Frameworks for feature combination
  - Deep learning based entity summarization
- What Lies Ahead

What is entity summarization?



What is entity summarization?



Google The Keyword Latest stories Product updates Company news

SEARCH

Introducing the Knowledge Graph: things, not strings

Amit Singhal
SVP, Engineering

Published May 16, 2012

Search is a lot about discovery—the basic human need to learn and broaden your horizons. But searching still requires a lot of hard work by you, the user. So today I'm really excited to launch the Knowledge Graph, which will help you discover new information quickly and easily.

The Knowledge Graph enhances Google Search in three main ways to start:

1. Find the right thing

2. Get the best summary

With the Knowledge Graph, Google can better understand your query, so we can summarize relevant content around that topic, including key facts you're likely to need for that particular thing. For example, if you're looking for Marie Curie, you'll see when she was born and died, but you'll also get details on her education and scientific discoveries:

3. Go deeper and broader

What is entity summarization?

Google The Keyword Latest stories Product updates Company news

SEARCH

Introducing the Knowledge Graph: things, not strings

Amit Singhal
SVP, Engineering

Published May 16, 2012

Search is a lot about discovery—the basic human need to learn and broaden your horizons. But searching still requires a lot of hard work by you, the user. So today I'm really excited to launch the Knowledge Graph, which will help you discover new information quickly and easily.

The Knowledge Graph enhances Google Search in three main ways to start:

1. Find the right thing
2. Get the best summary

With the Knowledge Graph, Google can better understand your query, so we can summarize relevant content around that topic, including key facts you're likely to need for that particular thing. For example, if you're looking for Marie Curie, you'll see when she was born and died, but you'll also get details on her education and scientific discoveries:

3. Go deeper and broader

Illinois

US State

Illinois is a midwestern state bordering Indiana in the east and the Mississippi River in the west. Nicknamed "the Prairie State," it's marked by farmland, forests, rolling hills and wetlands. Chicago, one of the largest cities in the U.S, is in the northeast on the shores of Lake Michigan. It's famous for its skyscrapers, such as sleek, 1,451-ft. Willis Tower and the neo-Gothic Tribune Tower. — Google

Capital: Springfield

Governor: J. B. Pritzker (Democratic Party)

Population: 12.67 million (2019)

Senators: Tammy Duckworth (Democratic Party), Dick Durbin (Democratic Party)

What is entity summarization?

- A generalized definition

Entity summarization. The problem of entity summarization is formulated as finding an optimal summary:

$$\text{find} \quad \arg\max_{S \subseteq \text{Desc}(e)} \texttt{score}(S|T), \quad \text{subject to } |S| \leq k, \tag{3}$$

where $\texttt{score}(S|T)$ is the quality score of summary S given T.

Liu et al., JoWS 2021

What is entity summarization?

- A generalized definition

Entity summarization. The problem of entity summarization is formulated as finding an optimal summary:

$$\text{find} \quad \arg\max_{S \subseteq \text{Desc}(e)} \texttt{score}(S|T), \quad \text{subject to } |S| \leq k, \tag{3}$$

where $\texttt{score}(S|T)$ is the quality score of summary S given T.

- A ranking-based definition

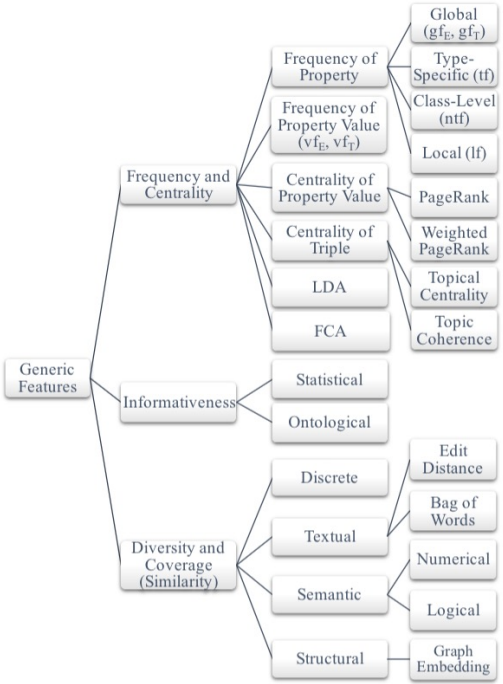
$$\texttt{score}(S|T) = \sum_{t \in S} \texttt{score}(t|T)$$

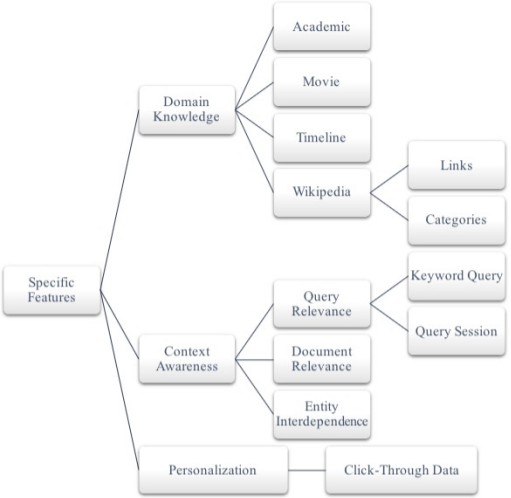
Liu et al., JoWS 2021

Outline

- Where We Are
  - What is entity summarization?
  - Technical features for entity summarization
  - Frameworks for feature combination
  - Deep learning based entity summarization
- What Lies Ahead

Technical features for entity summarization





Liu et al., JoWS 2021

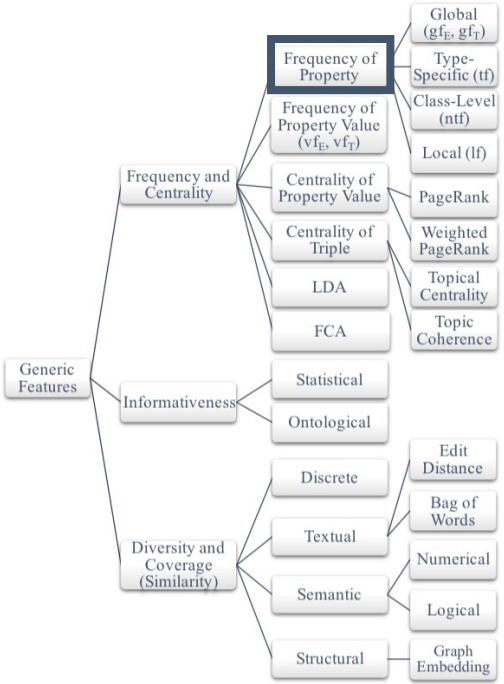
Technical features for entity summarization

Entity summarizers (sorted by publication date) and their technical features (Blanks: Not used).

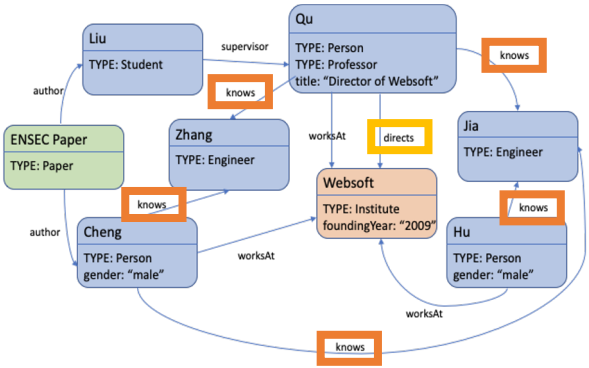
	Generic features			Specific features
	Frequency and centrality	Informativeness	Diversity and coverage (similarity)	Domains, contexts, and personalization
Falcons [32]			Bag-of-words	Query relevance
XRed [33]				Entity interdependence
Zhang et al. [34]	Weighted PageRank			Click-through data
RELIN [18]	Weighted PageRank	Statistical		
Thalhammer et al. [19]		Statistical		Movie domain
Yovisto [35]	tf			Academic domain, Wikipedia links
MMR-QSFS [36]				Query relevance
DIVERSUM [37]	lf		Discrete	
SUMMARUM [38]	PageRank		Discrete	
FACES [21]	$\mathtt{vf}_{\mathtt{T}}$	Statistical	Discrete, bag-of-words	
COMB [13]	lf	Statistical	Edit-distance-like, numerical, logical	Document relevance, entity interdependence
TimeMachine [39]				Timeline domain
C3D+P [15]	lf	Statistical	Edit-distance-like, numerical, logical	Entity interdependence
TRank^{++} [9]	$\mathtt{vf}_{\mathtt{E}}$	Ontological		Document relevance
FACES-E [22]	$\mathtt{vf}_{\mathtt{T}}$	Statistical	Discrete, bag-of-words	
CD [40]		Statistical	Edit-distance-like, numerical, logical	
Li et al. [23]	$\mathtt{gf}_{\mathtt{E}}$, lf			Movie domain
CES [20]	Weighted PageRank	Statistical		Session relevance
LinkSUM [24]	$\mathtt{gf}_{\mathtt{T}}$, lf, PageRank		Discrete	Wikipedia links
Aemoo [41]	tf			
DynES [42]	$\mathtt{gf}_{\mathtt{E}}$, $\mathtt{gf}_{\mathtt{T}}$, ntf, $\mathtt{vf}_{\mathtt{E}}$, $\mathtt{vf}_{\mathtt{T}}$			Query relevance
REMES [11]	$\mathtt{vf}_{\mathtt{E}}$	Statistical	Bag-of-words, structural	Entity interdependence
Multi-EGS [43]	tf, ntf, $\mathtt{vf}_{\mathtt{T}}$		Edit-distance-like	Wikipedia categories
ES-LDA [44]	LDA			Wikipedia categories
ES-LDA$_{ext}$ [45]	LDA			
CTab [16]	$\mathtt{gf}_{\mathtt{E}}$		Bag-of-words, numerical	Entity interdependence
BAFREC [46]	$\mathtt{gf}_{\mathtt{T}}$, $\mathtt{vf}_{\mathtt{T}}$	Ontological	Bag-of-words	
KAFCA [47]	FCA			
MPSUM [48]	LDA		Discrete	Wikipedia categories
Gottschalk et al. [49]				Timeline domain, Wikipedia links
VISION-KG [50]	$\mathtt{gf}_{\mathtt{E}}$, $\mathtt{gf}_{\mathtt{T}}$, ntf, $\mathtt{vf}_{\mathtt{E}}$, $\mathtt{vf}_{\mathtt{T}}$		Structural	Query relevance

Liu et al., JoWS 2021

Generic features --- frequency of property

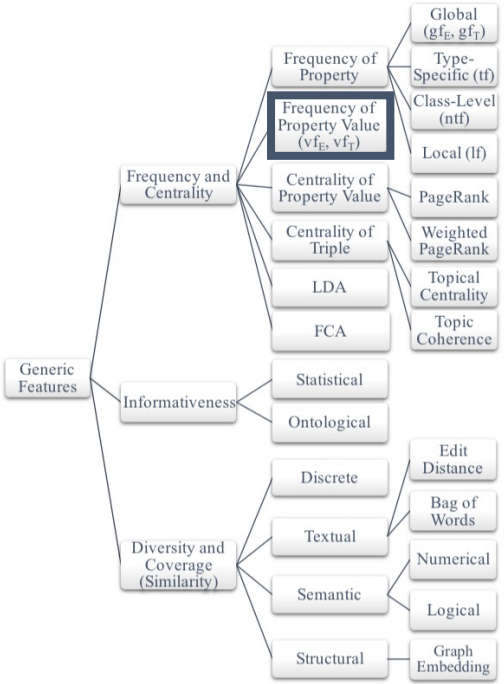


Frequency(knows) = 5
Frequency(directs) = 1

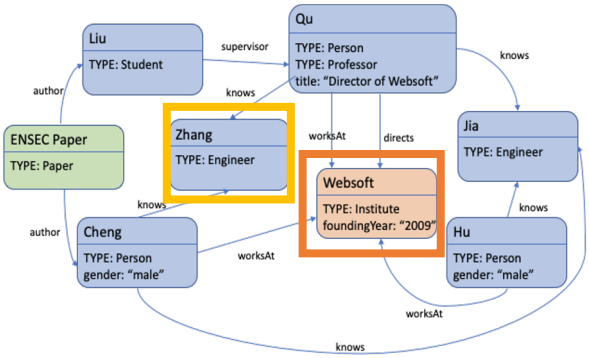


Thalhammer et al., ICWE 2016

Generic features --- frequency of property value

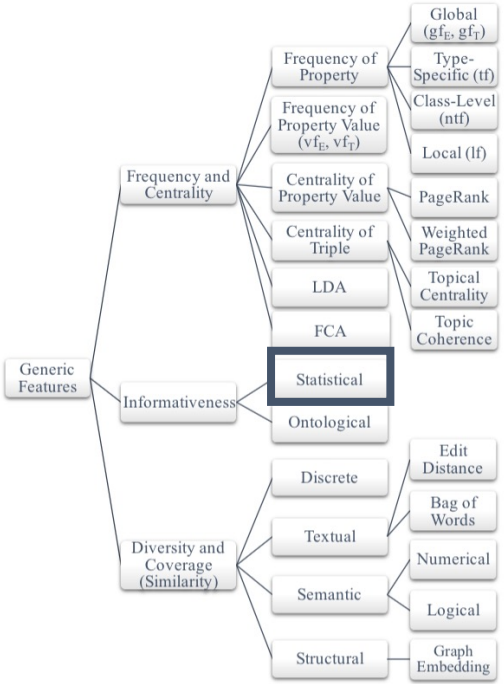


Frequency(Websoft) = 4
Frequency(Zhang) = 2

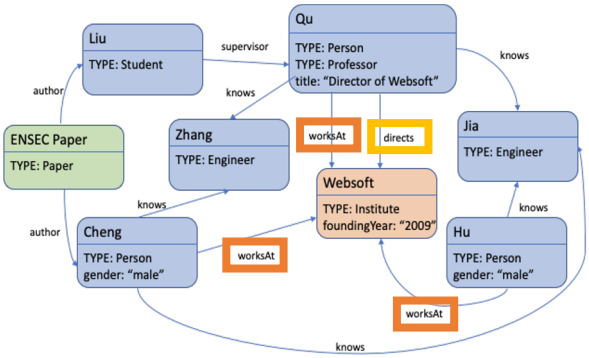


Gunaratna et al., AAAI 2015

Generic features --- statistical informativeness

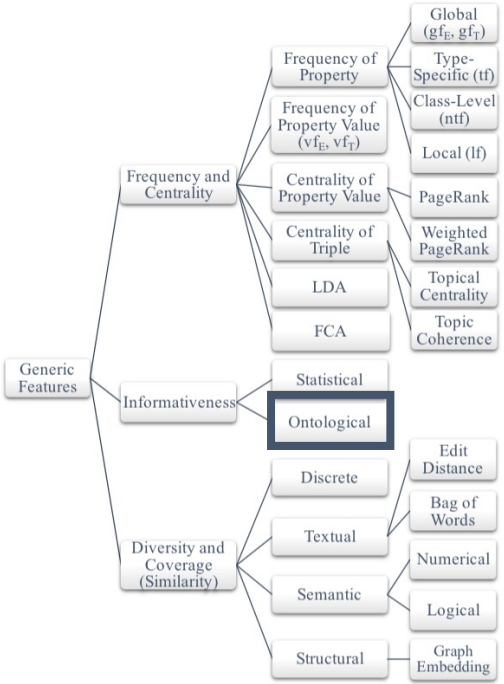


Info(worksAt, Websoft) = -log 3/8
Info(directs, Websoft) = -log 1/8

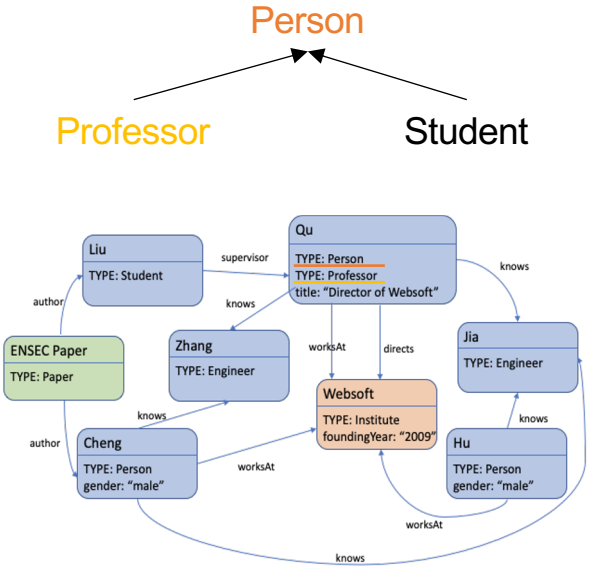


Cheng et al., ISWC 2011

Generic features --- ontological informativeness

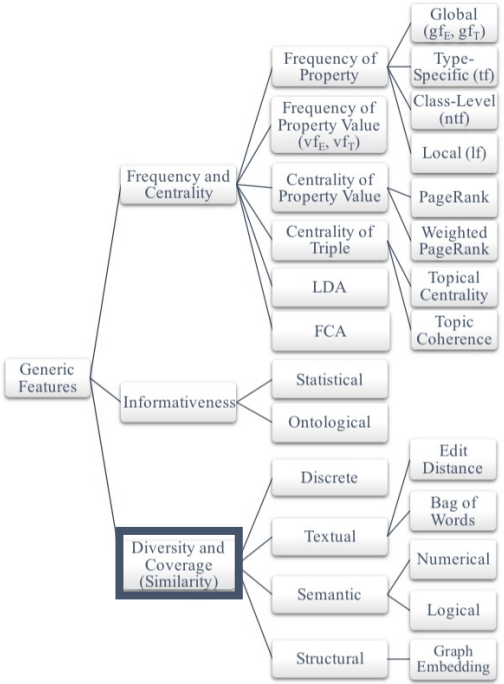


Info(Professor) > Info(Person)

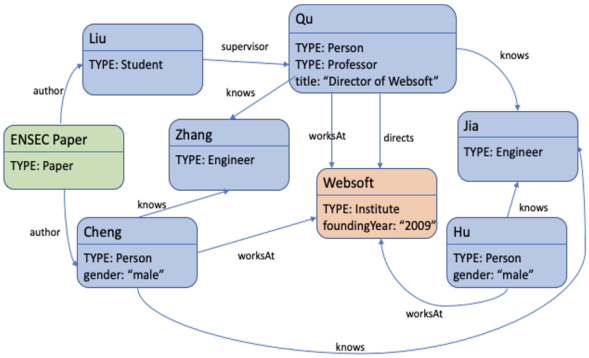


Tonon et al., JoWS 2016

Generic features --- diversity (similarity)

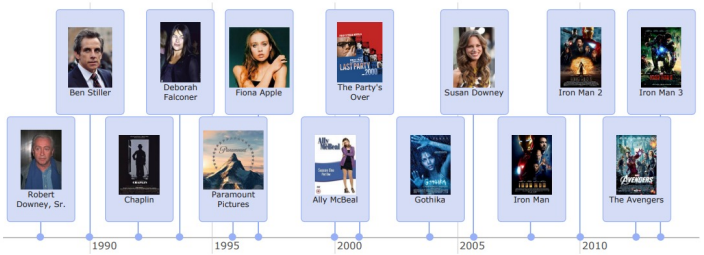


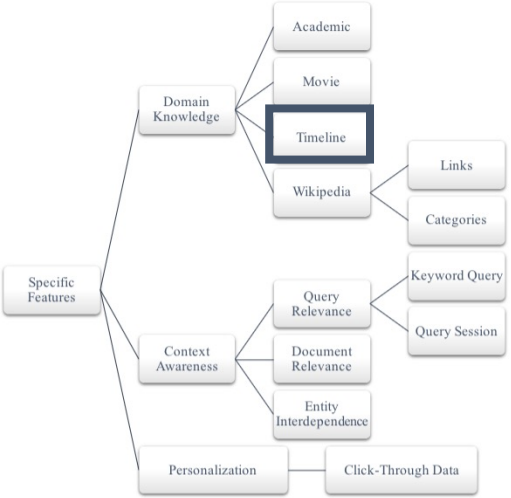
TextSim(Websoft, Director of Websoft)
NumSim(99, 100)
LogSim(TYPE-Person, TYPE-Professor)



Cheng et al., WWW 2015

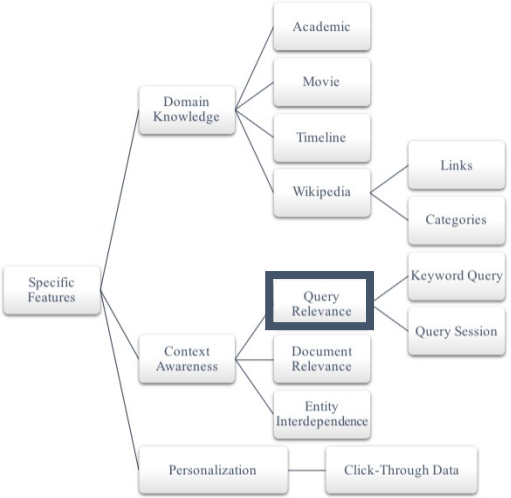
Specific features --- domain knowledge





Althoff et al., KDD 2015

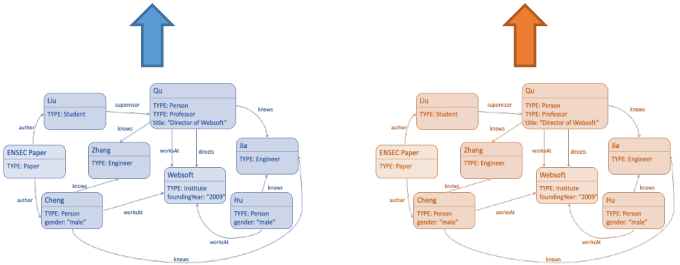
Specific features --- query relevance

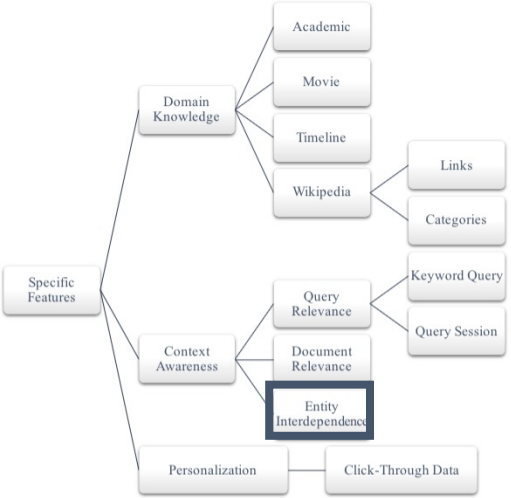


Cheng et al., IJSWIS 2009

Specific features --- entity interdependence

TimBL	*TBL*
⟨*givenname*, "Tim"⟩ ⟨*surname*, "Berners-Lee"⟩	⟨*name*, "Tim Berners-Lee"⟩
⟨*gender*, "male"⟩	⟨*sex*, "Male"⟩
⟨*is director of* , *W3C*⟩	⟨*invented*, *WWW*⟩ ⟨*founded*, *W3C*⟩





Cheng et al., JoWS 2015

Outline

- Where We Are
  - What is entity summarization?
  - Technical features for entity summarization
  - Frameworks for feature combination
  - Deep learning based entity summarization
- What Lies Ahead

Frameworks for feature combination

- simple frameworks
- random surfer model
- similarity-based grouping
- combinatorial optimization
- learning to rank

Frameworks for feature combination --- simple frameworks

- **Multiplication and summation**
  - Property ranking

    FRQ*EXC*DSC – relations are chosen by the product of frequency, exclusivity, and description.

  - Value ranking

$$score(e, r) = \alpha \cdot \frac{pr(r)}{max\{pr(a) : a \in res(e)\}} + (1 - \alpha) \cdot \mathbf{1}_{bl(e)}(r)$$

    (*pr* = PageRank, *bl* = backlink)

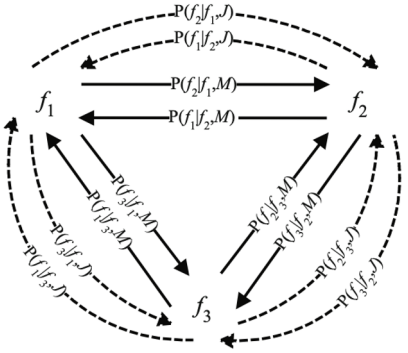
Pros: Be simple
Cons: Cannot represent diversity and coverage

Thalhammer et al., ICWE 2016

Frameworks for feature combination --- random surfer model

- **Random walk in a complete graph of property-value pairs**
  - Relational **M**ove
  - Informational **J**ump

$$\mathbf{x}_p(t+1) = \sum_{f_q \in \mathrm{FS}} \mathbf{x}_q(t) \cdot (\mathrm{P}(M|f_q) \cdot \mathrm{P}(f_p|f_q, M) + \mathrm{P}(J|f_q) \cdot \mathrm{P}(f_p|f_q, J))$$



$$\mathbf{M}_{p,q} = \sqrt{\mathrm{Rel}(\mathrm{Prop}(f_p), \mathrm{Prop}(f_q)) \cdot \mathrm{Rel}(\mathrm{Val}(f_p), \mathrm{Val}(f_q))}$$

$$\mathrm{PMI}(s_i, s_j) = \log \frac{\mathrm{P}(s_i, s_j)}{\mathrm{P}(s_i) \cdot \mathrm{P}(s_j)}$$

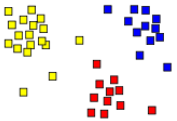
$$\mathbf{J}_{p,q} = \mathrm{SelfInfo}(f_p|f_q) = -\log(\mathrm{P}(f_p|f_q))$$

$$\mathrm{P}(f_p|f_q) = \frac{|\{e \in E \mid f_p, f_q \in \mathrm{FS}(e)\}|}{|\{e \in E \mid f_q \in \mathrm{FS}(e)\}|}$$

Pros: Be natural for combining centrality- and importance-style features
Cons: Cannot naturally represent diversity and coverage

Cheng et al., ISWC 2011

Frameworks for feature combination --- similarity-based grouping



- **Grouping property-value pairs by bag-of-words**
  - WordNet- and typing-based text expansion + Cobweb clustering

Feature (f)	Property expansion	Value expansion	Word set (WS(f))
birthPlace:Warsaw	{*birthPlace*, *birth*, *place*, beginning, point, area, locality, ..}	{*place*, *PopulatedPlace*, populated, point, area, locality, ..}	{*birthPlace*, *birth*, *place*, *PopulatedPlace*, beginning, populated, point, area, locality, ..}

  - Selecting top-ranked property-value pairs from as many clusters as possible
    - Ranking property-value pairs by self-information and value frequency

$$Rank(f) = Inf(f) * Po(Val(f)) \qquad Inf(f) = log(\frac{N}{|\{e|f \in FS(e)\}|})$$

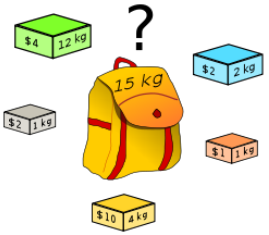
$$Po(v) = log|\{triple\ t|\exists\ e, f : t\ \text{“}appears\ in\text{”}\ G\ and\ t \equiv (e\ Prop(f)\ Val(f))\ and\ Val(f) = v\}|$$

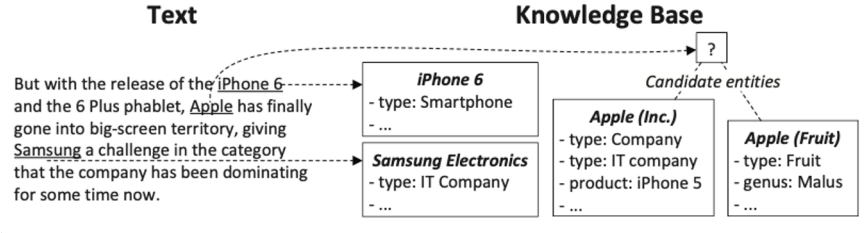
Pros: Can naturally represent diversity and coverage
Cons: Be inflexible --- essentially employing binary similarity

Gunaratna et al., AAAI 2015

Frameworks for feature combination --- combinatorial optimization

- **Quadratic multidimensional knapsack problem**
  - Objective function
    - Maximizing self-information and contextual relevance
    - Minimizing intra-entity property-value pair overlap
    - Maximizing inter-entity property-value pair difference
  - Constraints
    - Selecting at most *k* property-value pairs for each entity





Pros: Can flexibility employ numerical similarity
Cons: Be often NP-hard and sub-optimum

Cheng et al., WWW 2015

Frameworks for feature combination --- learning to rank

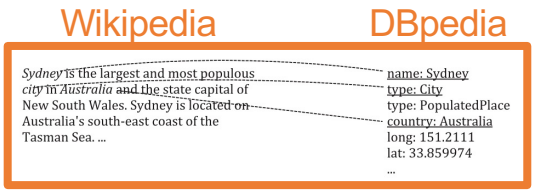
- Decision tree and linear regression
- Support vector machine
- Gradient tree boosting
- …

Pros: Automatically combine features
Cons: Need labeled data for training

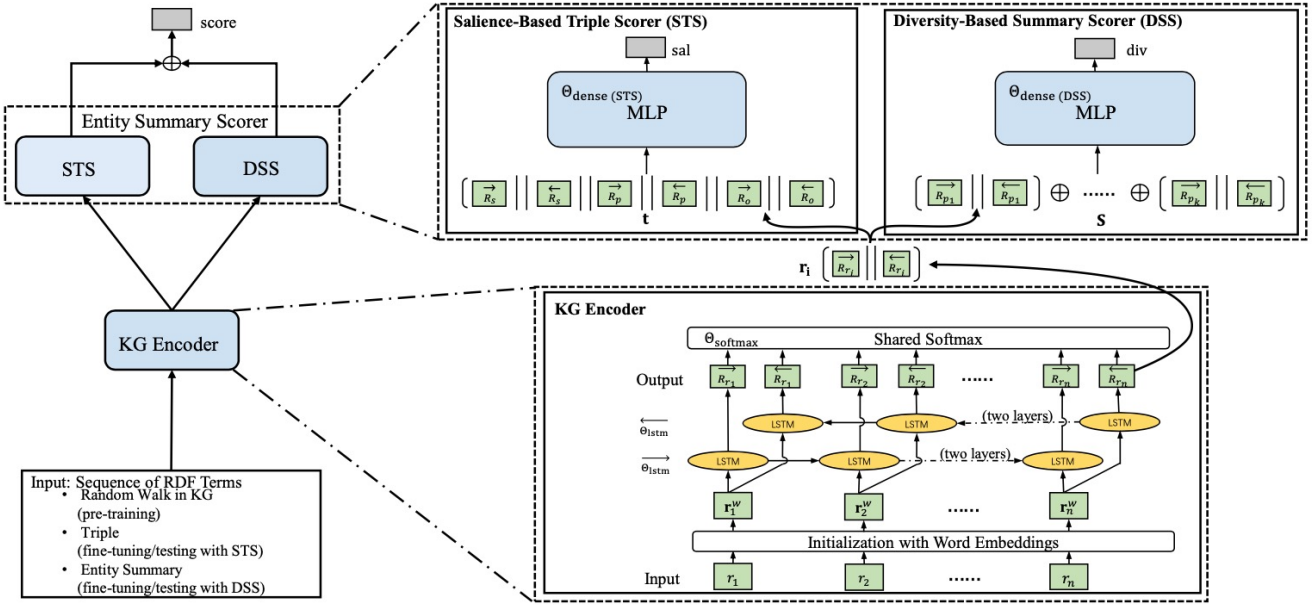
Outline

- Where We Are
  - What is entity summarization?
  - Technical features for entity summarization
  - Frameworks for feature combination
  - Deep learning based entity summarization
- What Lies Ahead

Deep learning based weakly supervised entity summarization



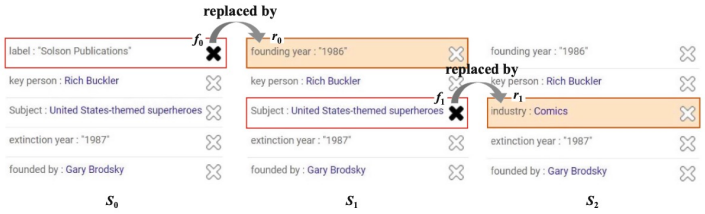
Weak supervision from programmatically labeled entity summaries

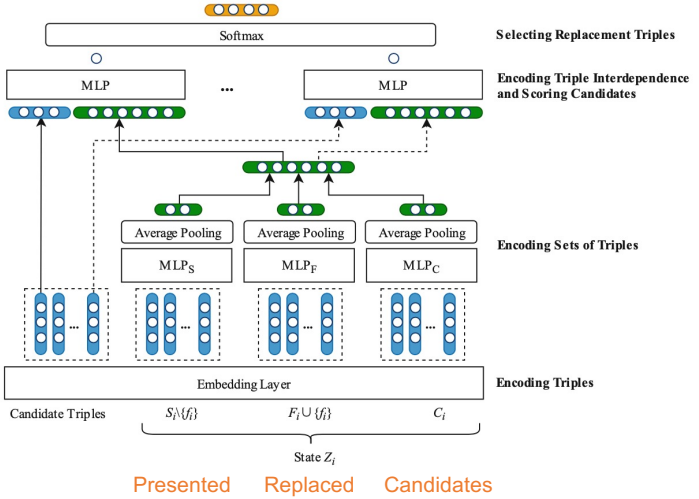


Li et al., IJCAI 2020

Deep learning based entity summarization with user feedback

- One summary does not fit all
- Iterative entity summarization as MDP





Liu et al., ESWC 2020 (best research paper award)

Outline

- Where We Are
  - What is entity summarization?
  - Technical features for entity summarization
  - Frameworks for feature combination
  - Deep learning based entity summarization
- What Lies Ahead

State-of-the-art results

- Mean F1

	DBpedia		LinkedMDB	
	$k = 5$	$k = 10$	$k = 5$	$k = 10$
RELIN	0.242 -○○▼▼▼▼▼▼	0.455 -▼○○▼○▼▼▼	0.203 -○○▼○▲▼○▼	0.258 -▼○▼▼○▼▼▼
DIVERSUM	0.249 ○-○○▼▼▼▼▼	0.507 ▲-▲○○○○○○	0.207 ○-○▼○▲▼○▼	0.358 ▲-▲○○▲▼○▼
FACES	0.270 ○○-○○○▼▼▼	0.428 ○▼-▼▼▼▼▼	0.169 ○○-▼▼○▼▼▼	0.263 ○▼-▼▼○▼▼▼
FACES-E	0.280 ▲○○-○○▼▼▼	0.488 ○○▲-○○○○○	0.313 ▲▲▲-▲▲▼▲○	0.393 ▲○▲-▲▲○○○
CD	0.283 ▲▲○○-○▼○○	0.513 ▲○▲○-○○○○	0.217 ○○▲▼-▲▼○▼	0.331 ▲○▲▼-▲▼▼▼
LinkSUM	0.287 ▲▲○○○-▼○○	0.486 ○○▲○○-○○○	0.140 ▼▼○▼▼-▼▼▼	0.279 ○▼○▼▼-▼▼▼
BAFREC	0.335 ▲▲▲▲▲▲-○○	0.503 ▲○▲○○○-○○	0.360 ▲▲▲▲▲▲-▲▲	0.402 ▲▲▲○▲▲-○○
KAFCA	0.314 ▲▲▲▲○○○-○	0.509 ▲○▲○○○○-○	0.244 ○○▲▼○▲▼-○	0.397 ▲○▲○▲▲○-○
MPSUM	0.314 ▲▲▲▲○○○○-	0.512 ▲○▲○○○○○-	0.272 ▲▲▲○▲▲▼○-	0.423 ▲▲▲○▲▲○○-
ORACLE	0.595	0.713	0.619	0.678

Liu et al., ESWC 2020 (best resource paper nominee)

Future directions

- Use of semantics
- Human factors
- Machine and deep learning
- Non-extractive methods
- Interactive methods
- …

Future directions --- use of semantics

- Textual and structural semantics
  - Deep representations?
- Ontological semantics
  - Axioms and logical reasoning?
- Their combination
  - Neural-symbolic methods?

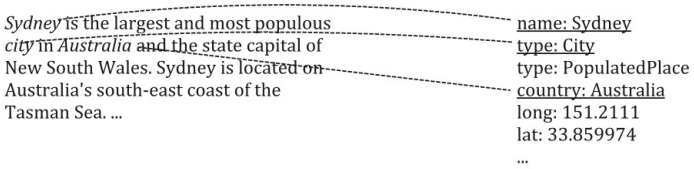
Future directions --- human factors

- Human friendliness
  - Readability?

Is ISBN an interesting property to human users?

Future directions --- machine and deep learning

- Labeled entity summaries for training
  - Weak supervision?
  - Crowd-sourcing?



Li et al., IJCAI 2020

Future directions --- non-extractive methods

- Form of a non-extractive (or abstractive) entity summary
  - Text?
  - Aggregation?

Triples	Atlas_Shrugged literaryGenre Science_fiction Atlas_Shrugged country United_States John_Galt series Atlas_Shrugged Atlas_Shrugged publicationYear ''1957'' Atlas_Shrugged author Ayn_Rand
Text Summary	Atlas Shrugged is a science fiction novel by Ayn Rand.

Vougiouklis et al., JAIR 2020

Future directions --- interactive methods

- Form of user feedback
  - Positive or negative opinions?
  - Explicit or implicit?
- Models and algorithms
  - Online learning to rank?

Liu et al., ESWC 2020 (best research paper award)

Future directions --- misc

- Personalization
- Contextualization
- Knowledge fusion
- …

References

- **Althoff et al., KDD 2015**, TimeMachine: Timeline generation for knowledge-base entities
- **Cheng et al., IJSWIS 2009**, Searching linked objects with falcons: Approach, implementation and evaluation
- **Cheng et al., ISWC 2011**, RELIN: relatedness and informativeness-based centrality for entity summarization
- **Cheng et al., WWW 2015**, Summarizing entity descriptions for effective and efficient human-centered entity linking
- **Cheng et al., JoWS 2015**, C3D+P: A summarization method for interactive entity resolution
- **Gunaratna et al., AAAI 2015**, FACES: diversity-aware entity summarization using incremental hierarchical conceptual clustering
- **Li et al., IJCAI 2020**, Neural entity summarization with joint encoding and weak supervision
- **Liu et al., ESWC 2020 (best research paper award)**, Entity summarization with user feedback
- **Liu et al., ESWC 2020 (best resource paper nominee)**, ESBM: an entity summarization benchmark
- **Liu et al., SSI 2020**, Entity summarization with high readability and low redundancy (in Chinese)
- **Liu et al., JoWS 2021**, Entity summarization: state of the art and future challenges
- **Thalhammer et al., ICWE 2016**, LinkSUM: using link analysis to summarize entity data
- **Tonon et al., JoWS 2016**, Contextualized ranking of entity types based on knowledge graphs
- **Vougiouklis et al., JAIR 2020**, Point at the triple: generation of text summaries from knowledge base triples

THANK
YOU