

# Knowledge Graphs: In Theory and Practice

**Nitish Aggarwal**, IBM Watson, USA,

**Sumit Bhatia**, IBM Research, India

Saeedeh Shekarpour, Knoesis Research Centre Ohio, USA

Amit Sheth, Knoesis Research Centre Ohio, USA

The material presented in this tutorial represents the personal opinion of the presenters and not of IBM and affiliated organization.

# Outline of the tutorial

## Part 1: Knowledge Graph Construction

- Introduction
- DBpedia: Knowledge extraction
- Approaches to extend knowledge graph
- Knowledge extraction from scratch

## Part 2: Knowledge Graph Analytics

- Finding entities of interest
- Entity exploration
- Upcoming challenges

# What is Knowledge Graph

*“The **Knowledge Graph** is a knowledge base used by Google to enhance its search engine's search results with semantic-search information gathered from a wide variety of sources.”*

# What is Knowledge Graph

*“The **Knowledge Graph** is a knowledge base used by Google to enhance its search engine's search results with semantic-search information gathered from a wide variety of sources.”*

*“A Knowledge graph (i) mainly describes real world entities and interrelations, organized in a graph (ii) defines possible classes and relations of entities in a schema” (iii) allows potentially interrelating arbitrary entities with each other... [Paulheim H.]*

*“We defines a Knowledge Graph as an RDF graph consists of a set of RDF triples where each RDF triple (s,p,o) is an ordered set of following RDF term ....” [Pujara J. al al.]*

# What is Knowledge Graph

No single formal definition ...

- Defines real world entities
- Provides relationships between them

# What is Knowledge Graph

No single formal definition ...

- Defines real world entities
- Provides relationships between them
- Contains rules defines through ontologies
- Enable reasoning to infer new knowledge

# Why Knowledge Graph

Building an intelligent system that can interact with human, requires knowledge about real world entities.

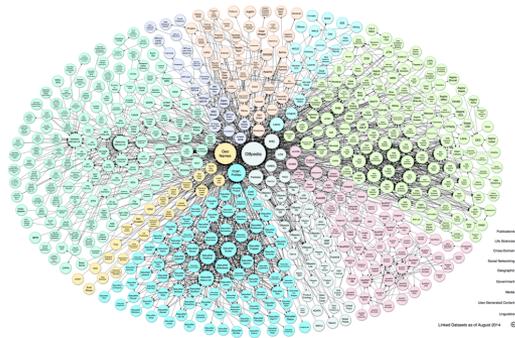
# Why Knowledge Graph

Building an intelligent system that can interact with human, requires knowledge about real world entities.

- Enhance search results.
- Enhance ad sense.
- Help in language understanding.
- Enables knowledge discovery.

Is there existing knowledge graph ready to use for my application?

LinkedIn  
Knowledge Graph



Amazon  
Product Graph

Facebook  
Entity Graph

Microsoft  
Satori





# DBpedia: Knowledge extraction

New York City

From Wikipedia, the free encyclopedia

Coordinates: 40°42′46″N 74°00′21″W

*"NYC" and "New York, New York" redirect here. For other uses, see [New York City \(disambiguation\)](#); [NYC \(disambiguation\)](#); and [New York, New York \(disambiguation\)](#).*

The **City of New York**, often called **New York City** or simply **New York**, is the most populous city in the United States.<sup>[9]</sup> With an estimated 2016 population of 8,537,673<sup>[7]</sup> distributed over a land area of about 302.6 square miles (784 km<sup>2</sup>),<sup>[10][11]</sup> New York City is also the most densely populated major city in the United States.<sup>[12]</sup> Located at the southern tip of the state of New York, the city is the center of the New York metropolitan area, one of the most populous urban agglomerations in the world,<sup>[13][14]</sup> with an estimated 23.7 million residents as of 2016.<sup>[5]</sup> A global power city,<sup>[15]</sup> New York City has been described as the cultural, financial, and media capital<sup>[16][17]</sup> of the world,<sup>[18][19][20][21][22]</sup> and exerts a significant impact upon commerce,<sup>[22]</sup> entertainment, research, technology, education, politics, and sports. The city's fast pace<sup>[23][24]</sup> defines the term *New York minute*.<sup>[25]</sup> Home to the headquarters of the United Nations,<sup>[26]</sup> New York is an important center for international diplomacy.<sup>[27]</sup>

Situated on one of the world's largest natural harbors,<sup>[28][29]</sup> New York City consists of five boroughs, each of which is a separate county of New York State.<sup>[30]</sup> The five boroughs – Brooklyn, Queens, Manhattan, The Bronx, and Staten Island – were consolidated into a single city in 1898.<sup>[31]</sup> The city and its metropolitan area constitute the premier gateway for legal immigration to the United States,<sup>[32]</sup> and as many as 800 languages are spoken in New York.<sup>[33][34][35]</sup> making it the most linguistically diverse city in the world.<sup>[34][36][37]</sup> New York City is home to more than 3.2 million residents born outside the United States,<sup>[38]</sup> the largest foreign-born population of any city in the world.<sup>[39]</sup> In 2013, the tri-state New York Metropolitan Statistical Area (MSA) produced a gross metropolitan product (GMP) of nearly US\$1.4 trillion.<sup>[40]</sup> If New York City were a country, it would have the 12th highest GDP in the world.<sup>[41]</sup>

New York City traces its origins to a trading post founded by colonists of the Dutch Republic in 1624 on Lower Manhattan; the post was named *New Amsterdam* in 1626.<sup>[42]</sup> The city and its surroundings came under English control in 1664<sup>[42]</sup> and were renamed *New York* after King Charles II of England granted the lands to his brother, the Duke of York.<sup>[43]</sup> New York served as the capital of the United States from 1785 until 1790.<sup>[44]</sup> It has been the country's largest city since 1790.<sup>[45]</sup> The Statue of Liberty greeted millions of immigrants as they came to the Americas by ship in the late 19th and early 20th centuries<sup>[46]</sup> and is a world symbol of the United States and its ideals of liberty and peace.<sup>[47]</sup> In the 21st century, New York has emerged as a global node of creativity and entrepreneurship,<sup>[48]</sup> social tolerance,<sup>[49]</sup> and environmental sustainability.<sup>[50][51]</sup> and as a symbol of freedom and cultural diversity.<sup>[52]</sup>

Many districts and landmarks in New York City are well known, and the city received a record 61 million tourists in 2016,<sup>[53]</sup> hosting three of the world's ten most visited tourist attractions in 2013.<sup>[54]</sup> Several sources have ranked New York the most photographed city in the world.<sup>[55][56]</sup> Times Square, iconic as the world's "heart"<sup>[57]</sup> and its "Crossroads",<sup>[58]</sup> is the brightly illuminated hub of the Broadway Theater District,<sup>[59]</sup> one of the world's busiest pedestrian intersections,<sup>[60][61]</sup> and a major center of the world's entertainment industry.<sup>[62]</sup>

The names of many of the city's bridges, skyscrapers,<sup>[63]</sup> and parks are known around the world. Anchored by Wall Street in the Financial District of Lower Manhattan, New York City has been called both the most economically powerful city and the leading financial center of the world.<sup>[22][64][65][66]</sup> and the city is home to the world's two largest stock exchanges by total market capitalization, the New York Stock Exchange and NASDAQ.<sup>[67][68]</sup> Manhattan's real estate market is among the most expensive in the world.<sup>[69][70]</sup> Manhattan's Chinatown incorporates the highest concentration of Chinese people in the Western Hemisphere,<sup>[71][72]</sup> with multiple signature Chinatowns developing across the city.<sup>[73][74]</sup> Providing continuous 24/7 service,<sup>[75]</sup> the New York City Subway is one of the most extensive metro systems worldwide, with 472 stations in operation.<sup>[76][77][78]</sup> Over 120 colleges and universities are located in New York City, including Columbia University, New York University, and Rockefeller University, which have been ranked among the top universities in the world.<sup>[79][80]</sup>

## Contents [hide]

- History
  - Eymology
  - Early history
  - Dutch rule
  - English rule
  - American Revolution
  - Nineteenth century
  - Modern history
- Geography
  - Cityscape
  - Architecture
  - Boroughs
  - Climate
  - Parks
    - 2.5.1 National parks
    - 2.5.2 State parks
    - 2.5.3 City parks
  - Military installations
- Demographics
  - Population density
  - Race and ethnicity
  - Sexual orientation and gender identity
    - 3.3.1 Transgender contribution
  - Religion
  - Income
- Economy
  - 4.1 City economic overview
  - 4.2 Wall Street
  - 4.3 Silicon Alley

<span>New York City</span>	
City	
City of New York	
<div><div><div><div><div><span></span></div></div></div><div><div><div><span></span></div></div><div><div><span></span></div></div></div><div><div><div><span></span></div></div><div><div><span></span></div></div></div><div><div><div><span></span></div></div><div><div><span></span></div></div></div><div><div><div><span></span></div></div><div><div><span></span></div></div></div></div></div> <div>Clockwise, from top: Midtown Manhattan, Times Square, the Unisphere in Queens, the Brooklyn Bridge, Lower Manhattan with One World Trade Center, Central Park, the headquarters of the United Nations, and the Statue of Liberty</div>	
<div><div><span></span><div>Flag</div></div><div><span></span><div>Seal</div></div></div>	
Nickname(s): <span><i>See <a href="#">Nicknames of New York City</a></i></span>	
<div><div><div><div><div><span></span><div><div><span><span></span></span></div><div><div><span></span></div></div></div></div><div></div><span>Location within the U.S. state of New York</span></div></div></div><div><div><div><div><span></span><div><div><span><span></span></span></div><div><div><span></span></div></div></div></div><div></div><span>New York City</span></div></div></div></div> <div><div><span>Show map of the US</span></div><div><span>Show map of New York</span></div><div><span>Show all</span></div></div>	

Location in the contiguous United States and New York

# DBpedia: Knowledge extraction

New York City  
From Wikipedia, the free encyclopedia

Coordinates: 40°42′46″N 74°00′21″W﻿ / ﻿40.712778°N 74.005833°W﻿ / 40.712778; -74.005833

*"NYC" and "New York, New York" redirect here. For other uses, see New York City (disambiguation); NYC (disambiguation); and New York, New York (disambiguation).*

**The City of New York**, often called **New York City** or simply **New York**, is the most populous city in the United States.<sup>[9]</sup> With an estimated 2016 population of 8,537,673<sup>[7]</sup> distributed over a land area of about 302.6 square miles (784 km<sup>2</sup>),<sup>[10][11]</sup> New York City is also the most densely populated major city in the United States.<sup>[12]</sup> Located at the southern tip of the state of New York, the city is the center of the New York metropolitan area, one of the most populous urban agglomerations in the world.<sup>[13][14]</sup> with an estimated 23.7 million residents as of 2016.<sup>[5]</sup> A *global power city*,<sup>[15]</sup> New York City has been described as the cultural, financial, and media capital<sup>[16][17]</sup> of the world,<sup>[18][19][20][21][22]</sup> and exerts a significant impact upon commerce,<sup>[22]</sup> entertainment, research, technology, education, politics, and sports. The city's fast pace<sup>[23][24]</sup> defines the term *New York minute*.<sup>[25]</sup> Home to the headquarters of the United Nations,<sup>[26]</sup> New York is an important center for international diplomacy.<sup>[27]</sup>

Situated on one of the world's largest natural harbors,<sup>[28][29]</sup> New York City consists of five boroughs, each of which is a separate county of New York State.<sup>[30]</sup> The five boroughs – Brooklyn, Queens, Manhattan, The Bronx, and Staten Island – were consolidated into a single city in 1898.<sup>[31]</sup> The city and its metropolitan area constitute the premier gateway for legal immigration to the United States,<sup>[32]</sup> and as many as 800 languages are spoken in New York,<sup>[33][34][35]</sup> making it the most linguistically diverse city in the world.<sup>[34][36][37]</sup> New York City is home to more than 3.2 million residents born outside the United States,<sup>[38]</sup> the largest foreign-born population of any city in the world.<sup>[39]</sup> In 2013, the tri-state New York Metropolitan Statistical Area (MSA) produced a gross metropolitan product (GMP) of nearly US\$1.4 trillion.<sup>[40]</sup> If New York City were a country, it would have the 12th highest GDP in the world.<sup>[41]</sup>

New York City traces its origins to a trading post founded by colonists of the Dutch Republic in 1624 on Lower Manhattan; the post was named *New Amsterdam* in 1626.<sup>[42]</sup> The city and its surroundings came under English control in 1664<sup>[42]</sup> and were renamed *New York* after King Charles II of England granted the lands to his brother, the Duke of York.<sup>[43]</sup> New York served as the capital of the United States from 1785 until 1790.<sup>[44]</sup> It has been the country's largest city since 1790.<sup>[45]</sup> The Statue of Liberty greeted millions of immigrants as they came to the Americas by ship in the late 19th and early 20th centuries<sup>[46]</sup> and is a world symbol of the United States and its ideals of liberty and peace.<sup>[47]</sup> In the 21st century, New York has emerged as a global node of creativity and entrepreneurship,<sup>[48]</sup> social tolerance,<sup>[49]</sup> and environmental sustainability.<sup>[50][51]</sup> and as a symbol of freedom and cultural diversity.<sup>[52]</sup>

Many districts and landmarks in New York City are well known, and the city received a record 61 million tourists in 2016,<sup>[53]</sup> hosting three of the world's ten most visited tourist attractions in 2013.<sup>[54]</sup> Several sources have ranked New York the most photographed city in the world.<sup>[55][56]</sup> Times Square, iconic as the world's "heart"<sup>[57]</sup> and its "Crossroads",<sup>[58]</sup> is the brightly illuminated hub of the Broadway Theater District,<sup>[59]</sup> one of the world's busiest pedestrian intersections,<sup>[60][61]</sup> and a major center of the world's entertainment industry.<sup>[62]</sup> The names of many of the city's bridges, skyscrapers,<sup>[63]</sup> and parks are known around the world. Anchored by Wall Street in the Financial District of Lower Manhattan, New York City has been called both the most economically powerful city and the leading financial center of the world,<sup>[22][64][65][66]</sup> and the city is home to the world's two largest stock exchanges by total market capitalization, the New York Stock Exchange and NASDAQ.<sup>[67][68]</sup> Manhattan's real estate market is among the most expensive in the world.<sup>[69][70]</sup> Manhattan's Chinatown incorporates the highest concentration of Chinese people in the Western Hemisphere,<sup>[71][72]</sup> with multiple signature Chinatowns developing across the city.<sup>[73][74]</sup> Providing continuous 24/7 service,<sup>[75]</sup> the New York City Subway is one of the most extensive metro systems worldwide, with 472 stations in operation.<sup>[76][77][78]</sup> Over 120 colleges and universities are located in New York City, including Columbia University, New York University, and Rockefeller University, which have been ranked among the top universities in the world.<sup>[79][80]</sup>

- 1 History
  - 1.1 Etyymology
  - 1.2 Early history
  - 1.3 Dutch rule
  - 1.4 English rule
  - 1.5 American Revolution
  - 1.6 Nineteenth century
  - 1.7 Modern history
- 2 Geography
  - 2.1 Cityscape
  - 2.2 Architecture
  - 2.3 Boroughs
  - 2.4 Climate
  - 2.5 Parks
    - 2.5.1 National parks
    - 2.5.2 State parks
    - 2.5.3 City parks
  - 2.6 Military installations
- 3 Demographics
  - 3.1 Population density
  - 3.2 Race and ethnicity
  - 3.3 Sexual orientation and gender identity
    - 3.3.1 Transgender contribution
  - 3.4 Religion
  - 3.5 Income
- 4 Economy
  - 4.1 City economic overview
  - 4.2 Wall Street
  - 4.3 Silicon Alley

**The City of New York, often called New York City or simply New York, is the most populous city in the United States.**

New York City	
<span></span>	<span></span>
City	
City of New York	
<span></span> <span></span> <span></span> <span></span> <span></span>	
<div><div><div><div><div><span></span></div><div>Clockwise, from top: Midtown Manhattan, Times Square, the Unisphere in Queens, the Brooklyn Bridge, Lower Manhattan with One World Trade Center, Central Park, the headquarters of the United Nations, and the Statue of Liberty</div></div></div></div></div>	
<span></span> <span></span>	<span></span> <span></span>
Flag	Seal
Nickname(s): <i>See Nicknames of New York City</i>	
<span></span>	
Location within the U.S. state of New York	
<span></span> New York City	
<span></span> Show map of the US	
<span></span> Show map of New York	
<span></span> Location in the contiguous United States and New York	
Coordinates: <span><span><span><span><span>40°42′46″N</span> <span>74°00′21″W</span></span></span><span><span>﻿</span> / <span>﻿</span></span><span><span>40.712778°N 74.005833°W</span><span><span>﻿</span> / <span>40.712778; -74.005833</span></span></span></span></span>	
<b>Country</b>	<span><span><span></span></span><span> </span></span> United States
<b>State</b>	<span><span><span></span></span><span> </span></span> New York
<b>Counties / (Boroughs)</b>	Bronx, Kings (Brooklyn), New York (Manhattan), Queens, Richmond (Staten Island)
<b>Historic colonies</b>	<span><span><span></span></span><span> </span></span> New Netherland <ul style="list-style-type: none"> <li><span><span><span></span></span><span> </span></span>Province of New York</li> </ul>
<b>Settled</b>	1624
<b>Consolidated</b>	1898
<b>Named for</b>	James, Duke of York
<b>Government</b> <sup>[2]</sup>	
<span> </span> • <b>Type</b>	Mayor–Council
<span> </span> • <b>Body</b>	New York City Council
<span> </span> • <b>Mayor</b>	Bill de Blasio (D)
<b>Area</b> <sup>[1]</sup>	
<span> </span> • <b>Total</b>	468.484 sq mi (1,213.37 km <sup>2</sup> )
<span> </span> • <b>Land</b>	302.643 sq mi (783.84 km <sup>2</sup> )
<span> </span> • <b>Water</b>	165.841 sq mi (429.53 km <sup>2</sup> )
<span> </span> • <b>Metro</b>	13,318 sq mi (34,490 km <sup>2</sup> )

# DBpedia: Knowledge extraction

New York City  
From Wikipedia, the free encyclopedia

Coordinates: 40°42′46″N 74°00′21″W﻿ / ﻿40.712777777778°N 74.005833333333°W﻿ / 40.712777777778; -74.005833333333

"*NYC*" and "*New York, New York*" redirect here. For other uses, see *New York City (disambiguation)*; *NYC (disambiguation)*; and *New York, New York (disambiguation)*.

The **City of New York**, often called **New York City** or simply **New York**, is the most populous city in the United States.<sup>[9]</sup> With an estimated 2016 population of 8,537,673<sup>[7]</sup> distributed over a land area of about 302.6 square miles (784 km<sup>2</sup>),<sup>[10][11]</sup> New York City is also the most densely populated major city in the United States.<sup>[12]</sup> Located at the southern tip of the state of New York, the city is the center of the New York metropolitan area, one of the most populous urban agglomerations in the world,<sup>[13][14]</sup> with an estimated 23.7 million residents as of 2016.<sup>[5]</sup> A *global power city*,<sup>[15]</sup> New York City has been described as the cultural, financial, and media capital<sup>[16][17]</sup> of the world,<sup>[18][19][20][21][22]</sup> and exerts a significant impact upon commerce,<sup>[22]</sup> entertainment, research, technology, education, politics, and sports. The city's fast pace<sup>[23][24]</sup> defines the term *New York minute*.<sup>[25]</sup> Home to the headquarters of the United Nations,<sup>[26]</sup> New York is an important center for international diplomacy.<sup>[27]</sup>

Situated on one of the world's largest natural harbors,<sup>[28][29]</sup> New York City consists of five boroughs, each of which is a separate county of New York State.<sup>[30]</sup> The five boroughs – Brooklyn, Queens, Manhattan, The Bronx, and Staten Island – were consolidated into a single city in 1898.<sup>[31]</sup> The city and its metropolitan area constitute the premier gateway for legal immigration to the United States,<sup>[32]</sup> and as many as 800 languages are spoken in New York,<sup>[33][34][35]</sup> making it the most linguistically diverse city in the world.<sup>[34][36][37]</sup> New York City is home to more than 3.2 million residents born outside the United States,<sup>[38]</sup> the largest foreign-born population of any city in the world.<sup>[39]</sup> In 2013, the tri-state New York Metropolitan Statistical Area (MSA) produced a gross metropolitan product (GMP) of nearly US\$1.4 trillion.<sup>[40]</sup> If New York City were a country, it would have the 12th highest GDP in the world.<sup>[41]</sup>

New York City traces its origins to a trading post founded by colonists of the Dutch Republic in 1624 on Lower Manhattan; the post was named *New Amsterdam* in 1626.<sup>[42]</sup> The city and its surroundings came under English control in 1664<sup>[42]</sup> and were renamed *New York* after King Charles II of England granted the lands to his brother, the Duke of York.<sup>[43]</sup> New York served as the capital of the United States from 1785 until 1790.<sup>[44]</sup> It has been the country's largest city since 1790.<sup>[45]</sup> The Statue of Liberty greeted millions of immigrants as they came to the Americas by ship in the late 19th and early 20th centuries<sup>[46]</sup> and is a world symbol of the United States and its ideals of liberty and peace.<sup>[47]</sup> In the 21st century, New York has emerged as a global node of creativity and entrepreneurship,<sup>[48]</sup> social tolerance,<sup>[49]</sup> and environmental sustainability.<sup>[50][51]</sup> and as a symbol of freedom and cultural diversity.<sup>[52]</sup>

Many districts and landmarks in New York City are well known, and the city received a record 61 million tourists in 2016,<sup>[53]</sup> hosting three of the world's ten most visited tourist attractions in 2013.<sup>[54]</sup> Several sources have ranked New York the most photographed city in the world.<sup>[55][56]</sup> Times Square, iconic as the world's "heart"<sup>[57]</sup> and its "Crossroads",<sup>[58]</sup> is the brightly illuminated hub of the Broadway Theater District,<sup>[59]</sup> one of the world's busiest pedestrian intersections,<sup>[60][61]</sup> and a major center of the world's entertainment industry.<sup>[62]</sup> The names of many of the city's bridges, skyscrapers,<sup>[63]</sup> and parks are known around the world. Anchored by Wall Street in the Financial District of Lower Manhattan, New York City has been called both the most economically powerful city and the leading financial center of the world,<sup>[22][64][65][66]</sup> and the city is home to the world's two largest stock exchanges by total market capitalization, the New York Stock Exchange and NASDAQ.<sup>[67][68]</sup> Manhattan's real estate market is among the most expensive in the world.<sup>[69][70]</sup> Manhattan's Chinatown incorporates the highest concentration of Chinese people in the Western Hemisphere,<sup>[71][72]</sup> with multiple signature Chinatowns developing across the city.<sup>[73][74]</sup> Providing continuous 24/7 service,<sup>[75]</sup> the New York City Subway is one of the most extensive metro systems worldwide, with 472 stations in operation.<sup>[76][77][78]</sup> Over 120 colleges and universities are located in New York City, including Columbia University, New York University, and Rockefeller University, which have been ranked among the top universities in the world.<sup>[79][80]</sup>

New York City	
City	
City of New York	
<div><div><div><div><div><span></span></div></div></div><div><div><div><span></span></div></div><div><div><span></span></div></div></div><div><div><div><span></span></div></div><div><div><span></span></div></div></div><div><div><div><span></span></div></div><div><div><span></span></div></div></div><div><div><div><span></span></div></div><div><div><span></span></div></div></div></div></div> <div>Clockwise, from top: Midtown Manhattan, Times Square, the Unisphere in Queens, the Brooklyn Bridge, Lower Manhattan with One World Trade Center, Central Park, the headquarters of the United Nations, and the Statue of Liberty</div>	
<div><div><div><span></span><div><div><span><span></span></span></div><div><span>Flag</span></div></div></div><div><div><span></span><div><div><span><span></span></span></div><div><span>Seal</span></div></div></div></div></div></div>	
Nickname(s): <i>See Nicknames of New York City</i>	
<div><div><div><span></span><div><div><span><span></span></span></div><div><span>Location within the U.S. state of New York</span></div></div></div><div><div><div><span></span><div><div><span><span></span></span></div><div><span>New York City</span></div></div></div><div><div><span>Show map of the US</span></div><div><span>Show map of New York</span></div></div><div><div><span>Show all</span></div></div></div></div>Location in the contiguous United States and New York</div></div>	
Coordinates: <span><span><span><span><span>40°42′46″N</span> <span>74°00′21″W</span></span></span><span><span>﻿</span> / <span>﻿</span></span><span><span>40.712777777778°N 74.005833333333°W</span><span><span>﻿</span> / <span>40.712777777778; -74.005833333333</span></span></span></span></span>	
Country	<span><span><span></span></span><span> </span></span> United States
State	<span><span><span></span></span><span> </span></span> New York
Counties / (Boroughs)	Bronx, Kings (Brooklyn), New York (Manhattan), Queens, Richmond (Staten Island)
Historic colonies	<span><span><span></span></span><span> </span></span> New Netherland <div><span><span><span></span></span><span> </span></span>Province of New York</div>
Settled	1624
Consolidated	1898
Named for	James, Duke of York
Government <sup>[2]</sup>	
<span> </span> • <span> </span> Type	Mayor–Council
<span> </span> • <span> </span> Body	New York City Council
<span> </span> • <span> </span> Mayor	Bill de Blasio (D)
Area <sup>[1]</sup>	
<span> </span> • <span> </span> Total	468.484 <span> </span> sq <span> </span> mi (1,213.37 <span> </span> km <sup>2</sup> )
<span> </span> • <span> </span> Land	302.643 <span> </span> sq <span> </span> mi (783.84 <span> </span> km <sup>2</sup> )
<span> </span> • <span> </span> Water	165.841 <span> </span> sq <span> </span> mi (429.53 <span> </span> km <sup>2</sup> )
<span> </span> • <span> </span> Metro	13,318 <span> </span> sq <span> </span> mi (34,490 <span> </span> km <sup>2</sup> )

- 1 History
  - 1.1 Etyymology
  - 1.2 Early history
  - 1.3 Dutch rule
  - 1.4 English rule
  - 1.5 American Revolution
  - 1.6 Nineteenth century
  - 1.7 Modern history
- 2 Geography
  - 2.1 Cityscape
  - 2.2 Architecture
  - 2.3 Boroughs
  - 2.4 Climate
  - 2.5 Parks
    - 2.5.1 National parks
    - 2.5.2 State parks
    - 2.5.3 City parks
  - 2.6 Military installations
- 3 Demographics
  - 3.1 Population density
  - 3.2 Race and ethnicity
  - 3.3 Sexual orientation and gender identity
    - 3.3.1 Transgender contribution
  - 3.4 Religion
  - 3.5 Income
- 4 Economy
  - 4.1 City economic overview
  - 4.2 Wall Street
  - 4.3 Silicon Alley

The City of New York, often called **New York City** or simply New York, is the most populous city in the **United States**.

<New York City>, <CityIn> <United States>.  
<City Name>, <locatedIn> <Country Name>.

# DBpedia: Knowledge extraction

## Berlin

From Wikipedia, the free encyclopedia

*This article is about the capital of Germany. For other uses, see [Berlin \(disambiguation\)](#).*

**Berlin** (/bɜːrˈlɪn, bɜːr-/, German: [bɛʁˈliːn] ( listen)) is the **capital** and the largest city of **Germany** as well as one of its 16 constituent **states**. With a population of approximately 3.7 million,<sup>[4]</sup> Berlin is the **second most populous city proper** in the **European Union** and the seventh **most populous urban area** in the European Union.<sup>[5]</sup> Located in northeastern Germany on the banks of the rivers **Spree** and **Havel**, it is the centre of the **Berlin-Brandenburg Metropolitan Region**, which has

Berlin is in  
gardens, ri

## Mumbai

From Wikipedia, the free encyclopedia

*"Bombay" redirects here. For other uses, see [Bombay \(disambiguation\)](#).*

*For other uses, see [Mumbai \(disambiguation\)](#).*

**Mumbai** (/mʊmˈbaɪ/; also known as **Bombay**, the official name until 1995) is the **capital city** of the Indian state of **Maharashtra**. It is the **most populous city in India** with an estimated city population of 18.4 million. Along with the neighbouring regions of the **Mumbai Metropolitan Region**, it is second most populous **metropolitan area in India**, with a population of 21.3 million as of 2016.<sup>[12][13][14]</sup> Mumbai lies on the **Konkan** on the west coast of India and has a deep **natural harbour**.

# DBpedia: Knowledge extraction

## IBM

From Wikipedia, the free encyclopedia  
(Redirected from *ibm*)

For other uses of *IBM*, see *IBM (disambiguation)*. "*Big Blue*" redirects here. It is not to be confused with *New York Giants*. For other uses of *Big Blue*, see *Big Blue (disambiguation)*.

**IBM (International Business Machines Corporation)** is an American multinational technology company headquartered in Armonk, New York, United States, with operations in over 170 countries. The company originated in 1911 as the *Computing-Tabulating-Recording Company* (CTR) and was renamed "International Business Machines" in 1924.

IBM manufactures and markets computers hardware, middleware and software, and offers hosting and consulting services in areas ranging from mainframe computers to nanotechnology. IBM is also a major research organization, holding the record for most patents generated by a business (as of 2017) for 24 consecutive years.<sup>[5]</sup> Inventions by IBM include the automated teller machine (ATM), the PC, the floppy disk, the hard disk drive, the magnetic stripe card, the relational database, the SQL programming language, the UPC barcode, and dynamic random-access memory (DRAM). The IBM mainframe, exemplified by the System/360, was the dominant computing platform during the 1960s and 1970s.

IBM has continually shifted its business mix by commoditizing markets focusing on higher-value, more profitable markets. This includes spinning off printer manufacturer Lexmark in 1991 and selling off its personal computer (ThinkPad/ThinkCentre) and x86-based server businesses to Lenovo (2005 and 2014, respectively), and acquiring companies such as PwC Consulting (2002), SPSS (2009), and The Weather Company (2016). Also in 2014, IBM announced that it would go "fabless", continuing to design semiconductors, but offloading manufacturing to GlobalFoundries.

Nicknamed **Big Blue**, IBM is one of 30 companies included in the Dow Jones Industrial Average and one of the world's largest employers, with (as of 2016) nearly 380,000 employees. Known as "IBMers", IBM employees have been awarded five Nobel Prizes, six Turing Awards, ten National Medals of Technology and five National Medals of Science.

### Contents [hide]

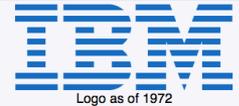
- History
- Headquarters and offices
- Products and services
- Research
- Brand and reputation
- People and culture
  - Employees
    - IBM alumni
  - Board and shareholders
- See also
- References
- Further reading
- External links

## History [edit]

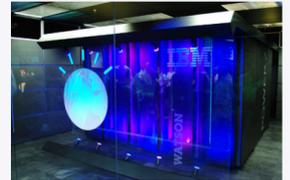
*Main article: History of IBM*

In the 1880s, technologies emerged that would ultimately form the core of International Business Machines (IBM). Julius E. Pitrat patented the computing scale in 1885;<sup>[6]</sup> Alexander Dey invented the dial recorder (1888);<sup>[7]</sup> Herman Hollerith patented the *Electric Tabulating Machine*;<sup>[8]</sup> and Willard Bundy invented a time clock to record a worker's arrival and departure time on a paper tape in 1889.<sup>[9]</sup> On June 16, 1911, their four companies were amalgamated in New York State by Charles Ranlett Flint forming a fifth company, the *Computing-Tabulating-Recording Company* (CTR) based in Endicott, New York.<sup>[1][10]</sup> The five companies had 1,300 employees and offices and plants in Endicott and Binghamton, New York; Dayton, Ohio; Detroit, Michigan; Washington, D.C.; and Toronto. They manufactured machinery for sale and lease, ranging from commercial scales and industrial time recorders, meat and cheese slicers, to tabulators and punched cards. Thomas J. Watson, Sr., fired from the *National Cash Register Company* by John Henry Patterson, called on Flint and, in 1914, was offered CTR.<sup>[11]</sup> Watson joined CTR as *General Manager* then, 11 months later, was made *President* when court cases relating to his time at NCR were resolved.<sup>[12]</sup> Having learned Patterson's pioneering business practices, Watson proceeded to put the stamp of NCR onto CTR's companies.<sup>[13]</sup> He implemented sales conventions, "generous sales incentives, a focus on customer service, an insistence on well-groomed, dark-suited salesmen and had an evangelical fervor for instilling company pride and loyalty in every worker".<sup>[14][15]</sup> His favorite slogan, "THINK", became a mantra for each company's employees.<sup>[14]</sup> During Watson's first four years, revenues reached \$9 million and the company's operations expanded to Europe, South America, Asia and Australia.<sup>[14]</sup> "Watson had never liked the clumsy hyphenated title of the CTR" and in 1924 chose to replace it with the more expansive

### International Business Machines Corporation



Logo as of 1972



IBM Watson system in 2011

<b>Type</b>	Public
<b>Traded as</b>	NYSE: <span>IBM</span> <span>↗</span> <div>DJIA Component</div> <div>S&amp;P 100 Component</div> <div>S&amp;P 500 Component</div>
<b>ISIN</b>	US4592001014
<b>Industry</b>	Cloud computing · Cognitive computing
<b>Founded</b>	June 16, 1911; 106 years ago (as Computing-Tabulating-Recording Company) <div>Endicott, New York, U.S.<sup>[1]</sup></div> <div>Charles Ranlett Flint</div>
<b>Founder</b>	Armonk, New York, U.S.
<b>Area served</b>	177 countries <sup>[2]</sup>
<b>Key people</b>	Ginni Rometty (Chairwoman, President and CEO)
<b>Products</b>	See IBM products
<b>Revenue</b>	<span>▼</span> US\$ 79.919 billion (2016) <sup>[3]</sup>
<b>Operating income</b>	<span>▼</span> US\$ 13.031 billion (2016) <sup>[3]</sup>
<b>Net income</b>	<span>▼</span> US\$ 11.872 billion (2016) <sup>[3]</sup>
<b>Total assets</b>	<span>▲</span> US\$ 117.47 billion (2016) <sup>[3]</sup>
<b>Total equity</b>	<span>▲</span> US\$ 18.392 billion (2016) <sup>[3]</sup>
<b>Number of employees</b>	380,000 (2016) <sup>[4]</sup>
<b>Website</b>	<span>www.ibm.com</span> <span>↗</span>

# DBpedia: Knowledge extraction

## IBM

From Wikipedia, the free encyclopedia  
(Redirected from *ibm*)

For other uses of IBM, see *IBM (disambiguation)*. "*Big Blue*" redirects here. It is not to be confused with *New York Giants*. For other uses of *Big Blue*, see *Big Blue (disambiguation)*.

**IBM (International Business Machines Corporation)** is an American multinational technology company headquartered in Armonk, New York, United States, with operations in over 170 countries. The company originated in 1911 as the *Computing-Tabulating-Recording Company* (CTR) and was renamed "International Business Machines" in 1924.

IBM manufactures and markets computers hardware, middleware and software, and offers hosting and consulting services in areas ranging from mainframe computers to nanotechnology. IBM is also a major research organization, holding the record for most patents generated by a business (as of 2017) for 24 consecutive years.<sup>[5]</sup> Inventions by IBM include the automated teller machine (ATM), the PC, the floppy disk, the hard disk drive, the magnetic stripe card, the relational database, the SQL programming language, the UPC barcode, and dynamic random-access memory (DRAM). The IBM mainframe, exemplified by the System/360, was the dominant computing platform during the 1960s and 1970s.

IBM has continually shifted its business mix by commoditizing markets focusing on higher-value, more profitable markets. This includes spinning off printer manufacturer Lexmark in 1991 and selling off its personal computer (ThinkPad/ThinkCentre) and x86-based server businesses to Lenovo (2005 and 2014, respectively), and acquiring companies such as PwC Consulting (2002), SPSS (2009), and The Weather Company (2016). Also in 2014 IBM announced that it would go "fabless", continuing to design semiconductors, but offloading manufacturing to GlobalFoundries.

Nicknamed **Big Blue**, IBM is one of 30 companies included in the Dow Jones Industrial Average and one of the world's largest employers, with (as of 2016) nearly 380,000 employees. Known as "IBMers", IBM employees have been awarded five Nobel Prizes, six Turing Awards, ten National Medals of Technology and five National Medals of Science.

### Contents [hide]

- History
- Headquarters and offices
- Products and services
- Research
- Brand and reputation
- People and culture
  - Employees
    - IBM alumni
    - Board and shareholders
- See also
- References
- Further reading
- External links

<head entity>, <rel> <tail entity >

## History [edit]

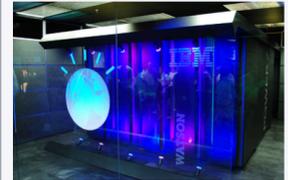
*Main article: History of IBM*

In the 1880s, technologies emerged that would ultimately form the core of International Business Machines (IBM). Julius E. Pitrat patented the computing scale in 1885;<sup>[6]</sup> Alexander Dey invented the dial recorder (1888);<sup>[7]</sup> Herman Hollerith patented the *Electric Tabulating Machine*;<sup>[8]</sup> and Willard Bundy invented a time clock to record a worker's arrival and departure time on a paper tape in 1889.<sup>[9]</sup> On June 16, 1911, their four companies were amalgamated in New York State by Charles Ranlett Flint forming a fifth company, the *Computing-Tabulating-Recording Company* (CTR) based in Endicott, New York.<sup>[11][10]</sup> The five companies had 1,300 employees and offices and plants in Endicott and Binghamton, New York; Dayton, Ohio; Detroit, Michigan; Washington, D.C.; and Toronto. They manufactured machinery for sale and lease, ranging from commercial scales and industrial time recorders, meat and cheese slicers, to tabulators and punched cards. Thomas J. Watson, Sr., fired from the *National Cash Register Company* by John Henry Patterson, called on Flint and, in 1914, was offered CTR.<sup>[11]</sup> Watson joined CTR as *General Manager* then, 11 months later, was made *President* when court cases relating to his time at NCR were resolved.<sup>[12]</sup> Having learned Patterson's pioneering business practices, Watson proceeded to put the stamp of NCR onto CTR's companies.<sup>[13]</sup> He implemented sales conventions, "generous sales incentives, a focus on customer service, an insistence on well-groomed, dark-suited salesmen and had an evangelical fervor for instilling company pride and loyalty in every worker".<sup>[14][15]</sup> His favorite slogan, "THINK", became a mantra for each company's employees.<sup>[14]</sup> During Watson's first four years, revenues reached \$9 million and the company's operations expanded to Europe, South America, Asia and Australia.<sup>[14]</sup> "Watson had never liked the clumsy hyphenated title of the CTR" and in 1924 chose to replace it with the more expansive

### International Business Machines Corporation



Logo as of 1972



IBM Watson system in 2011

<b>Type</b>	Public
<b>Traded as</b>	NYSE: <span>IBM</span> <span>📄</span> <div>DJIA Component S&amp;P 100 Component S&amp;P 500 Component</div>
<b>ISIN</b>	US4592001014
<b>Industry</b>	Cloud computing · Cognitive computing
<b>Founded</b>	June 16, 1911; 106 years ago (as Computing-Tabulating-Recording Company) Endicott, New York, U.S. <sup>[1]</sup> Charles Ranlett Flint
<b>Founder</b>	Charles Ranlett Flint
<b>Headquarters</b>	Armonk, New York, U.S.
<b>Area served</b>	177 countries <sup>[2]</sup>
<b>Key people</b>	Ginni Rometty (Chairwoman, President and CEO)
<b>Products</b>	See IBM products
<b>Revenue</b>	▼ US\$ 79.919 billion (2016) <sup>[3]</sup>
<b>Operating income</b>	▼ US\$ 13.031 billion (2016) <sup>[3]</sup>
<b>Net income</b>	▼ US\$ 11.872 billion (2016) <sup>[3]</sup>
<b>Total assets</b>	▲ US\$ 117.47 billion (2016) <sup>[3]</sup>
<b>Total equity</b>	▲ US\$ 18.392 billion (2016) <sup>[3]</sup>
<b>Number of employees</b>	380,000 (2016) <sup>[4]</sup>
<b>Website</b>	<span>www.ibm.com</span> <span>🌐</span>

# DBpedia: Knowledge extraction

## IBM

From Wikipedia, the free encyclopedia  
(Redirected from *ibm*)

For other uses of IBM, see *IBM (disambiguation)*. "*Big Blue*" redirects here. It is not to be confused with *New York Giants*. For other uses of *Big Blue*, see *Big Blue (disambiguation)*.

**IBM (International Business Machines Corporation)** is an American multinational technology company headquartered in Armonk, New York, United States, with operations in over 170 countries. The company originated in 1911 as the *Computing-Tabulating-Recording Company* (CTR) and was renamed "International Business Machines" in 1924.

IBM manufactures and markets computers hardware, middleware and software, and offers hosting and consulting services in areas ranging from mainframe computers to nanotechnology. IBM is also a major research organization, holding the record for most patents generated by a business (as of 2017) for 24 consecutive years.<sup>[5]</sup> Inventions by IBM include the automated teller machine (ATM), the PC, the floppy disk, the hard disk drive, the magnetic stripe card, the relational database, the SQL programming language, the UPC barcode, and dynamic random-access memory (DRAM). The IBM mainframe, exemplified by the System/360, was the dominant computing platform during the 1960s and 1970s.

IBM has continually shifted its business mix by commoditizing markets focusing on higher-value, more profitable markets. This includes spinning off printer manufacturer Lexmark in 1991 and selling off its personal computer (ThinkPad/ThinkCentre) and x86-based server businesses to Lenovo (2005 and 2014, respectively), and acquiring companies such as PwC Consulting (2002), SPSS (2009), and The Weather Company (2016). Also in 2014 IBM announced that it would go "fabless", continuing to design semiconductors, but offloading manufacturing to GlobalFoundries.

Nicknamed **Big Blue**, IBM is one of 30 companies included in the Dow Jones Industrial Average and one of the world's largest employers, with (as of 2016) nearly 380,000 employees. Known as "IBMers", IBM employees have been awarded five Nobel Prizes, six Turing Awards, ten National Medals of Technology and five National Medals of Science.

### Contents [hide]

- 1 History
- 2 Headquarters and offices
- 3 Products and services
- 4 Research
- 5 Brand and reputation
- 6 People and culture
  - 6.1 Employees
    - 6.1.1 IBM alumni
  - 6.2 Board and shareholders
- 7 See also
- 8 References
- 9 Further reading
- 10 External links

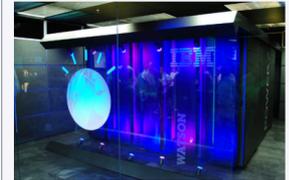
<head entity>, <rel> <tail entity >

## Wikipedia Infobox

### International Business Machines Corporation



Logo as of 1972



IBM Watson system in 2011

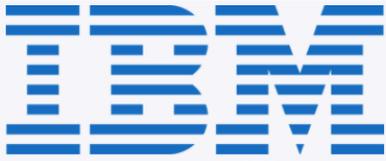
<b>Type</b>	Public
<b>Traded as</b>	NYSE: <span>IBM</span> <span>↗</span> <div>DJIA Component</div> <div>S&amp;P 100 Component</div> <div>S&amp;P 500 Component</div>
<b>ISIN</b>	US4592001014
<b>Industry</b>	Cloud computing · Cognitive computing
<b>Founded</b>	June 16, 1911; 106 years ago (as Computing-Tabulating-Recording Company) <div>Endicott, New York, U.S.<sup>[1]</sup></div> Charles Ranlett Flint
<b>Founder</b>	Charles Ranlett Flint
<b>Headquarters</b>	Armonk, New York, U.S.
<b>Area served</b>	177 countries <sup>[2]</sup>
<b>Key people</b>	Ginni Rometty (Chairwoman, President and CEO)
<b>Products</b>	See IBM products
<b>Revenue</b>	<span>▼</span> US\$ 79.919 billion (2016) <sup>[3]</sup>
<b>Operating income</b>	<span>▼</span> US\$ 13.031 billion (2016) <sup>[3]</sup>
<b>Net income</b>	<span>▼</span> US\$ 11.872 billion (2016) <sup>[3]</sup>
<b>Total assets</b>	<span>▲</span> US\$ 117.47 billion (2016) <sup>[3]</sup>
<b>Total equity</b>	<span>▲</span> US\$ 18.392 billion (2016) <sup>[3]</sup>
<b>Number of employees</b>	380,000 (2016) <sup>[4]</sup>
<b>Website</b>	<span>www.ibm.com</span> <span>↗</span>

## History [edit]

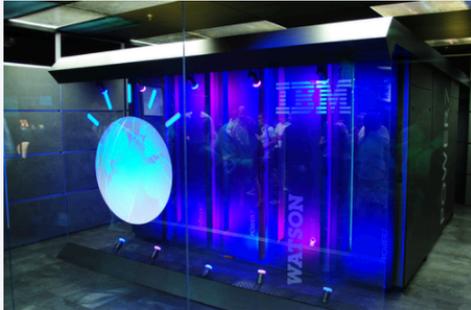
*Main article: History of IBM*

In the 1880s, technologies emerged that would ultimately form the core of International Business Machines (IBM). Julius E. Pitrat patented the computing scale in 1885;<sup>[6]</sup> Alexander Dey invented the dial recorder (1888);<sup>[7]</sup> Herman Hollerith patented the *Electric Tabulating Machine*;<sup>[8]</sup> and Willard Bundy invented a time clock to record a worker's arrival and departure time on a paper tape in 1889.<sup>[9]</sup> On June 16, 1911, their four companies were amalgamated in New York State by Charles Ranlett Flint forming a fifth company, the *Computing-Tabulating-Recording Company* (CTR) based in Endicott, New York.<sup>[11][10]</sup> The five companies had 1,300 employees and offices and plants in Endicott and Binghamton, New York; Dayton, Ohio; Detroit, Michigan; Washington, D.C.; and Toronto. They manufactured machinery for sale and lease, ranging from commercial scales and industrial time recorders, meat and cheese slicers, to tabulators and punched cards. Thomas J. Watson, Sr., fired from the *National Cash Register Company* by John Henry Patterson, called on Flint and, in 1914, was offered CTR.<sup>[11]</sup> Watson joined CTR as *General Manager* then, 11 months later, was made *President* when court cases relating to his time at NCR were resolved.<sup>[12]</sup> Having learned Patterson's pioneering business practices, Watson proceeded to put the stamp of NCR onto CTR's companies.<sup>[13]</sup> He implemented sales conventions, "generous sales incentives, a focus on customer service, an insistence on well-groomed, dark-suited salesmen and had an evangelical fervor for instilling company pride and loyalty in every worker".<sup>[14][15]</sup> His favorite slogan, "THINK", became a mantra for each company's employees.<sup>[14]</sup> During Watson's first four years, revenues reached \$9 million and the company's operations expanded to Europe, South America, Asia and Australia.<sup>[14]</sup> "Watson had never liked the clumsy hyphenated title of the CTR" and in 1924 chose to replace it with the more expansive

**International Business Machines Corporation**



Logo as of 1972



IBM Watson system in 2011

<b>Type</b>	Public
<b>Traded as</b>	NYSE: <a href="#">IBM</a> DJIA Component S&P 100 Component S&P 500 Component
<b>ISIN</b>	US4592001014
<b>Industry</b>	Cloud computing · Cognitive computing
<b>Founded</b>	June 16, 1911; 106 years ago (as Computing-Tabulating-Recording Company) Endicott, New York, U.S. <sup>[1]</sup>
<b>Founder</b>	Charles Ranlett Flint
<b>Headquarters</b>	Armonk, New York, U.S.
<b>Area served</b>	177 countries <sup>[2]</sup>
<b>Key people</b>	Ginni Rometty (Chairwoman, President and CEO)
<b>Products</b>	See <a href="#">IBM products</a>
<b>Revenue</b>	▼ US\$ 79.919 billion (2016) <sup>[3]</sup>
<b>Operating income</b>	▼ US\$ 13.031 billion (2016) <sup>[3]</sup>
<b>Net income</b>	▼ US\$ 11.872 billion (2016) <sup>[3]</sup>
<b>Total assets</b>	▲ US\$ 117.47 billion (2016) <sup>[3]</sup>
<b>Total equity</b>	▲ US\$ 18.392 billion (2016) <sup>[3]</sup>

**Apple Inc.**



The Apple Campus in Cupertino, California

<b>Formerly called</b>	Apple Computer Company (1976–1977) Apple Computer, Inc. (1977–2007)
<b>Type</b>	Public
<b>Traded as</b>	NASDAQ: <a href="#">AAPL</a> NASDAQ-100 component DJIA component S&P 100 component S&P 500 component
<b>ISIN</b>	US0378331005
<b>Industry</b>	Computer hardware Computer software Consumer electronics Digital distribution Semiconductors Fabless silicon design Corporate venture capital
<b>Founded</b>	April 1, 1976; 41 years ago
<b>Founders</b>	Steve Jobs Steve Wozniak Ronald Wayne
<b>Headquarters</b>	Apple Campus, 1 Infinite Loop, Cupertino, California, U.S.

**Google LLC**



Google's logo since 2015



Google's headquarters, the Googleplex, in August 2014

<b>Type</b>	Subsidiary
<b>Industry</b>	Internet Software Computer hardware
<b>Founded</b>	September 4, 1998; 19 years ago in Menlo Park, California, U.S. <sup>[1][2]</sup>
<b>Founders</b>	Larry Page Sergey Brin
<b>Headquarters</b>	Googleplex, Mountain View, California, U.S. <sup>[3]</sup>
<b>Area served</b>	Worldwide
<b>Key people</b>	Sundar Pichai (CEO) Ruth Porat (CFO)
<b>Products</b>	List of Google products
<b>Number of employees</b>	57,100 <sup>[4]</sup> (2015)
<b>Parent</b>	Alphabet Inc. (2015–present)
<b>Subsidiaries</b>	List of subsidiaries
<b>Website</b>	<a href="#">google.com</a>

# DBpedia: Knowledge extraction

## Ontology Classes

- owl:Thing
  - Activity (edit)
    - Game (edit)
      - BoardGame (edit)
      - CardGame (edit)
    - Sales (edit)
    - Sport (edit)
      - Athletics (edit)
      - Boxing (edit)
        - BoxingCategory (edit)
        - BoxingStyle (edit)
      - HorseRiding (edit)
      - TeamSport (edit)
        - Soccer (edit)
  - Agent (edit)
    - Deity (edit)
    - Employer (edit)
    - Family (edit)
      - NobleFamily (edit)
    - FictionalCharacter (edit)
      - ComicsCharacter (edit)
        - AnimangaCharacter (edit)
      - DisneyCharacter (edit)
      - MythologicalFigure (edit)
      - NarutoCharacter (edit)
      - SoapCharacter (edit)
    - Organisation (edit)
      - Broadcaster (edit)
        - BroadcastNetwork (edit)
        - RadioStation (edit)
        - TelevisionStation (edit)

- Organisation (edit)
  - Broadcaster (edit)
    - BroadcastNetwork (edit)
    - RadioStation (edit)
    - TelevisionStation (edit)
  - Company (edit)
    - Bank (edit)
    - Brewery (edit)
    - Caterer (edit)
    - LawFirm (edit)
    - PublicTransitSystem (edit)
      - Airline (edit)
      - BusCompany (edit)
    - Publisher (edit)
    - RecordLabel (edit)
    - Winery (edit)
  - EducationalInstitution (edit)
    - College (edit)
    - Library (edit)
    - School (edit)
    - University (edit)
  - EmployersOrganisation (edit)
  - GeopoliticalOrganisation (edit)
  - GovernmentAgency (edit)
    - GovernmentCabinet (edit)
  - Group (edit)
    - Band (edit)
    - ComedyGroup (edit)

- Person (edit)
  - Archeologist (edit)
  - Architect (edit)
  - Aristocrat (edit)
  - Artist (edit)
    - Actor (edit)
      - AdultActor (edit)
      - VoiceActor (edit)
    - Comedian (edit)
    - ComicsCreator (edit)
    - Dancer (edit)
    - FashionDesigner (edit)
    - Humorist (edit)
    - MusicalArtist (edit)
      - BackScene (edit)
      - ClassicalMusicArtist (edit)
      - Instrumentalist (edit)
        - Guitarist (edit)
      - MusicDirector (edit)
      - Singer (edit)
    - Painter (edit)
    - Photographer (edit)
    - Sculptor (edit)
  - Astronaut (edit)
  - Athlete (edit)
    - ArcherPlayer (edit)
    - AthleticsPlayer (edit)

# DBpedia: Knowledge extraction

## Ontology Classes

- owl:Thing
  - Activity (edit)
    - Game (edit)
      - BoardGame (edit)
      - CardGame (edit)
    - Sales (edit)
    - Sport (edit)
      - Athletics (edit)
      - Boxing (edit)
        - BoxingCategory (edit)
        - BoxingStyle (edit)
      - HorseRiding (edit)
      - TeamSport (edit)
        - Soccer (edit)
  - Agent (edit)
    - Deity (edit)
    - Employer (edit)
    - Family (edit)
      - NobleFamily (edit)
    - FictionalCharacter (edit)
      - ComicsCharacter (edit)
        - AnimangaCharacter (edit)
      - DisneyCharacter (edit)
      - MythologicalFigure (edit)
      - NarutoCharacter (edit)
      - SoapCharacter (edit)
    - Organisation (edit)
      - Broadcaster (edit)
        - BroadcastNetwork (edit)
        - RadioStation (edit)
        - TelevisionStation (edit)

- Organisation (edit)
  - Broadcaster (edit)
    - BroadcastNetwork (edit)
    - RadioStation (edit)
    - TelevisionStation (edit)
  - Company (edit)
    - Bank (edit)
    - Brewery (edit)
    - Caterer (edit)
    - LawFirm (edit)
    - PublicTransitSystem (edit)
      - Airline (edit)
      - BusCompany (edit)
    - Publisher (edit)
    - RecordLabel (edit)
    - Winery (edit)
  - EducationalInstitution (edit)
    - College (edit)
    - Library (edit)
    - School (edit)
    - University (edit)
  - EmployersOrganisation (edit)
  - GeopoliticalOrganisation (edit)
  - GovernmentAgency (edit)
    - GovernmentCabinet (edit)
  - Group (edit)
    - Band (edit)
    - ComedyGroup (edit)

- Person (edit)
  - Archeologist (edit)
  - Architect (edit)
  - Aristocrat (edit)
  - Artist (edit)
    - Actor (edit)
      - AdultActor (edit)
      - VoiceActor (edit)
    - Comedian (edit)
    - ComicsCreator (edit)
    - Dancer (edit)
    - FashionDesigner (edit)
    - Humorist (edit)
    - MusicalArtist (edit)
      - BackScene (edit)
      - ClassicalMusicArtist (edit)
      - Instrumentalist (edit)
        - Guitarist (edit)
      - MusicDirector (edit)
      - Singer (edit)
    - Painter (edit)
    - Photographer (edit)
    - Sculptor (edit)
  - Astronaut (edit)
  - Athlete (edit)
    - ArcherPlayer (edit)
    - AthleticsPlayer (edit)

# DBpedia: Knowledge extraction

## Properties on Actor:

Name	Label	Domain	Range
academyAward ( <a href="#">edit</a> )	Academy Award	<a href="#">Artist</a>	<a href="#">Award</a>
afiAward ( <a href="#">edit</a> )	AFI Award	<a href="#">Artist</a>	<a href="#">Award</a>
arielAward ( <a href="#">edit</a> )	Ariel Award	<a href="#">Actor</a>	<a href="#">Award</a>
associatedAct ( <a href="#">edit</a> )	associated act	<a href="#">Artist</a>	<a href="#">Artist</a>
baftaAward ( <a href="#">edit</a> )	BAFTA Award	<a href="#">Artist</a>	<a href="#">Award</a>
cesarAward ( <a href="#">edit</a> )	Cesar Award	<a href="#">Artist</a>	<a href="#">Award</a>
disciple ( <a href="#">edit</a> )	disciple	<a href="#">Artist</a>	<a href="#">Artist</a>
dutchRKDCoDe ( <a href="#">edit</a> )	Dutch RKD code	<a href="#">Artist</a>	<i>xsd:string</i>
emmyAward ( <a href="#">edit</a> )	Emmy Award	<a href="#">Artist</a>	<a href="#">Award</a>
field ( <a href="#">edit</a> )	field	<a href="#">Artist</a>	<a href="#">owl:Thing</a>
filmFareAward ( <a href="#">edit</a> )	Film Fare Award	<a href="#">Artist</a>	<a href="#">Award</a>
gaudiAward ( <a href="#">edit</a> )	Gaudí Award	<a href="#">Artist</a>	<a href="#">Award</a>
geminiAward ( <a href="#">edit</a> )	Gemini Award	<a href="#">Actor</a>	<a href="#">Award</a>
goldenCalfAward ( <a href="#">edit</a> )	Golden Calf Award	<a href="#">Actor</a>	<a href="#">Award</a>
goldenGlobeAward ( <a href="#">edit</a> )	Golden Globe Award	<a href="#">Artist</a>	<a href="#">Award</a>
goldenRaspberryAward ( <a href="#">edit</a> )	Golden Raspberry Award	<a href="#">Actor</a>	<a href="#">Award</a>
goyaAward ( <a href="#">edit</a> )	Goya Award	<a href="#">Artist</a>	<a href="#">Award</a>
grammyAward ( <a href="#">edit</a> )	Grammy Award	<a href="#">Artist</a>	<a href="#">Award</a>
iftaAward ( <a href="#">edit</a> )	IFTA Award	<a href="#">Actor</a>	<a href="#">Award</a>

# DBpedia: Knowledge extraction

Ontology  
(Classes, properties)

Parsers

dbr:IBM dbp:foundedBy  
dbr:Charles\_Ranlett\_Flint

dbr:IBM dbp:foundedBy  
dbr:Charles\_Ranlett\_Flint

dbr:IBM dbp:foundedBy  
dbr:Charles\_Ranlett\_Flint

.....

The screenshot shows the Wikipedia article for IBM. On the right side, a DBpedia parser interface is visible, displaying a table of extracted data. The table includes fields such as Name, Type, Founded, and Headquarters, with corresponding values extracted from the article text. The table is titled 'DBpedia: IBM' and has a 'View' button at the bottom.

Property	Value
Name	IBM
Type	Company
Founded	1911
Headquarters	Armonk, New York, US
Area served	Worldwide
Products	IBM Business Partner Program, IBM Cloud, IBM Consulting, IBM Financial Services, IBM Global Business Services, IBM Hybrid Cloud, IBM Quantum, IBM Research, IBM Security, IBM Software Group, IBM Systems, IBM Technology, IBM Z
Revenue	US\$62.5 billion (2022)
Employees	345,000 (2022)
Website	ibm.com

# (Research) problems in knowledge graphs

- Incomplete knowledge
  - Missing entities
  - Missing relations
  - Limited entity and relation types

# (Research) problems in knowledge graphs

- **Incomplete knowledge**
  - Missing entities
  - Missing relations
  - Limited entity and relation types
- **Incorrect knowledge**
  - Wrong entity label recognition
  - Wrong entity and relation type
  - Wrong facts

# (Research) problems in knowledge graphs

- **Incomplete knowledge**
  - Missing entities
  - Missing relations
  - Limited entity and relation types
- **Incorrect knowledge**
  - Wrong entity label recognition
  - Wrong entity and relation type
  - Wrong facts
- **Inconsistency in knowledge**
  - Different labels for same entity
  - Merging entities with same labels

# Approaches to extend knowledge graphs

- Extracting knowledge from Wikipedia tables
  - Large amount of raw data in form of tables
  - Tables have some implicit structure/patterns

# Approaches to extend knowledge graphs

- Extracting knowledge from Wikipedia tables
  - Large amount of raw data in form of tables
  - Tables have some implicit structure/patterns

No.	Position	Player	No.	Position	Player
1	 GK	André Onana	20	 MF	Lasse Schöne ( <i>vice-captain</i> )
3	 DF	Joël Veltman ( <i>captain</i> )	21	 MF	Frenkie de Jong
4	 DF	Matthijs de Ligt	22	 GK	Benjamin van Leer
5	 DF	Maximilian Wöber	23	 MF	Siem de Jong
6	 MF	Donny van de Beek	25	 FW	Kasper Dolberg
7	 FW	David Neres	26	 DF	Nick Viergever
8	 DF	Daley Sinkgraven	28	 DF	Luis Manuel Orejuela
9	 FW	Klaas-Jan Huntelaar	31	 GK	Norbert Alblas
10	 MF	Hakim Ziyech ( <i>3rd captain</i> )	33	 GK	Kostas Lamprou
11	 FW	Amin Younes	34	 MF	Abdelhak Nouri <sup>[110]</sup>
15	 MF	Carel Eiting	35	 DF	Mitchell Dijks
17	 FW	Václav Černý	42	 DF	Deyovaisio Zeefuik
19	 FW	Mateo Cassierra	45	 FW	Justin Kluivert

Wiki:AFC\_Ajax containing relations between players, their shirt number, and country

# Approaches to extend knowledge graphs

No.	Position	Player	No.	Position	Player
1	 GK	André Onana	20	 MF	Lasse Schöne ( <i>vice-captain</i> )
3	 DF	Joël Veltman ( <i>captain</i> )	21	 MF	Frenkie de Jong
4	 DF	Matthijs de Ligt	22	 GK	Benjamin van Leer
5	 DF	Maximilian Wöber	23	 MF	Siem de Jong
6	 MF	Donny van de Beek	25	 FW	Kasper Dolberg
7	 FW	David Neres	26	 DF	Nick Viergever
8	 DF	Daley Sinkgraven	28	 DF	Luis Manuel Orejuela
9	 FW	Klaas-Jan Huntelaar	31	 GK	Norbert Alblas
10	 MF	Hakim Ziyech ( <i>3rd captain</i> )	33	 GK	Kostas Lamprou
11	 FW	Amin Younes	34	 MF	Abdelhak Nouri <sup>[110]</sup>
15	 MF	Carel Eiting	35	 DF	Mitchell Dijks
17	 FW	Václav Černý	42	 DF	Deyovaisio Zeefuik
19	 FW	Mateo Cassierra	45	 FW	Justin Kluivert

- `<Wiki:AFC_Ajax, dbp:rel, Wiki:Andre_Onana>`
- 80% entities in the table have relation `dbp:rel` with the Wikipedia title entity `Wiki_AFC_Ajax`
- Other 20% entities are likely to have the same relationship `dbp:rel` with `Wiki_AFC_Ajax`

# Approaches to extend knowledge graphs

- Features
  - Article features: no. of tables, length
  - Table features: no. of rows, no. of columns
  - Column features: no. of entities in column, potential relations
  - Cell features: no. of entities in a cell, length of cell
  - Many others
- Combines using classification method

	<b>Prec.</b>	<b>Rec.</b>	<b>F1</b>
Rule-based	64.23	70.46	67.20
SVM	72.43	75.77	74.06
Logistic	79.62	79.01	79.31

# Approaches to extend knowledge graphs

- Features

- Article features: no. of tables, length
- Table features: no. of rows, no. of columns

- Rules/heuristics based methods makes mistakes, and hard to create one rule for everyone.
- Even though combining different features achieves 80% accuracy, it introduces 20% noise.

	<b>Prec.</b>	<b>Rec.</b>	<b>F1</b>
Rule-based	64.23	70.46	67.20
SVM	72.43	75.77	74.06
Logistic	79.62	79.01	79.31

Table data is limited, we need to go beyond

# Approaches to extend knowledge graphs

- Missing entity/literal for a relation
  - “Christopher A. Welty is an American computer scientist, who works at Google Research in NY”
    - `<dbr:Chris_Welty> <employedBy> <?>`
  - "Tom Cruise and Brad Pitt appear in Interview with the Vampire"
    - `<dbr:Brad_Pitt> <?> <dbr:Tom_Cruise>`

# Approaches to extend knowledge graphs

- Missing entity/literal for a relation
  - “Christopher A. Welty is an American computer scientist, who works at Google Research in NY”
    - `<dbr:Chris_Welty> <employedBy> <?>`
  - "Tom Cruise and Brad Pitt appear in Interview with the Vampire"
    - `<dbr:Brad_Pitt> <?> <dbr:Tom_Cruise>`
- Knowledge Base Completion
  - Similar to link prediction in social network but a bit more challenging
  - Need to identify relation type in addition to binary output.

# Approaches to extend knowledge graphs

- Knowledge Base Completion
  - TransE: learn the entity and relation embeddings by assuming that translation of entity embeddings correspond to their relation embeddings.  
[Bordes et al. 2013]
  - $S + R \approx T$ , where  $\langle S, R, T \rangle$
  - TransH: Learn different entity embedding for different relationships [Wang et al. 2014]
  - TransR: Learn entity and relation embeddings in different space, following by translation perform in relation space.  
[Lin Y. et al. 2015]
  - Many more methods [Nickel M. et al, 2015]

Knowledge base completion approaches focus on finding missing entities/reasons

Need to add new entities from external sources

## Need to add new entities from external sources

- Entity recognition in external text resource
  - Many Named Entity Recognition systems
- Link extracted entity to KG or create a new node if it does not have a corresponding entity
  - TAC-KBP (Entity Discovery and Linking task) [Ji H. et al. 2016]

Building knowledge graph such as DBpedia requires lot of manual efforts.

Building knowledge graph such as DBpedia requires lot of manual efforts.

- Many applications require domain/data specific custom knowledge graphs.
- Creating schema with class structure and constraints for each KG is difficult.

How to create a knowledge graph from unstructured text?

Jonathon Watson works at IBM. He has more than 50 patents, and won best inventor award for his invention “Neural Chip by Jon Watson et al.

Jonathon Watson works at IBM. He has more than 50 patents, and won best inventor award for his invention "Neural Chip by Jon Watson et al.

Entity  
extraction

Jonathon Watson  
IBM  
Jon Watson

Relation  
extraction

employedBy(Jonathon Watson,  
IBM)  
Jon Watson

Noise  
reduction

Jonathon  
Watson,  
Jon Watson

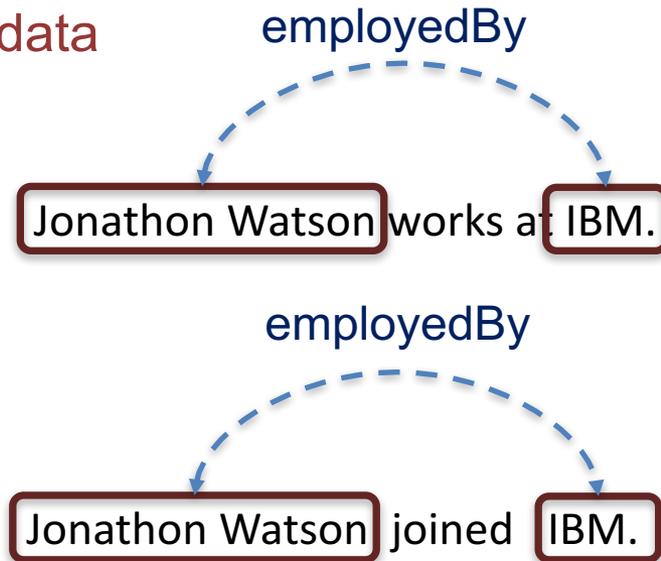


# Relation extraction

- Supervised methods

Predefined schema (employedBy, bornOn, BirthPlace ...)

Training data

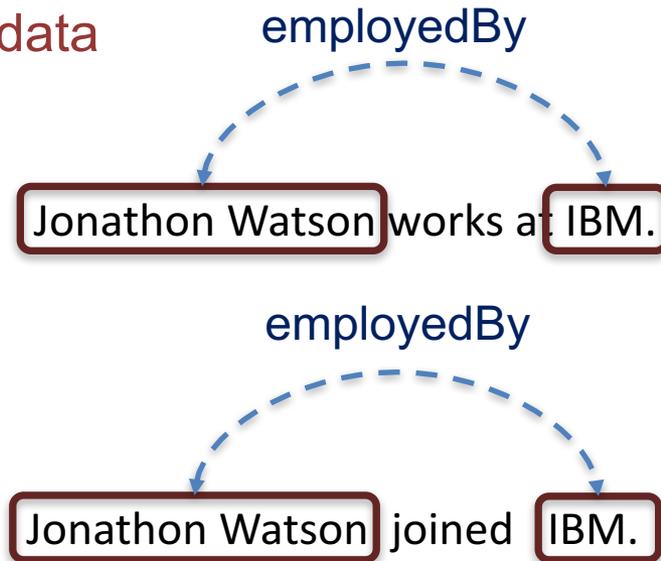


# Relation extraction

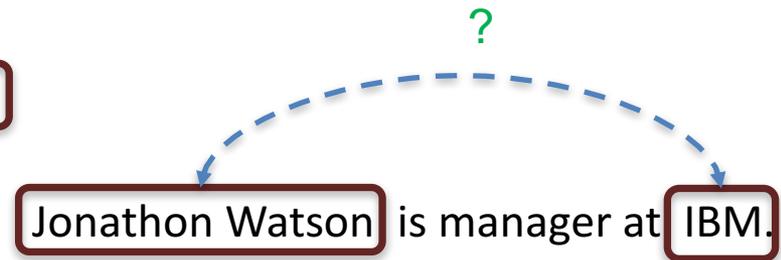
- Supervised methods

Predefined schema (employedBy, bornOn, BirthPlace ...)

Training data



Test data

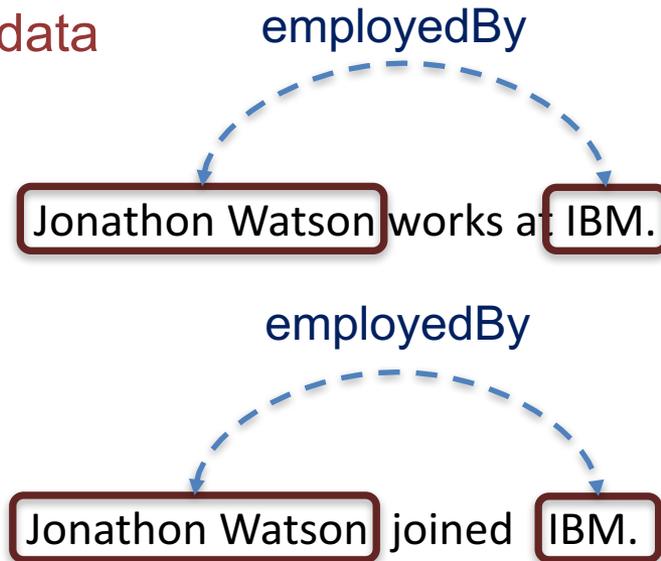


# Relation extraction

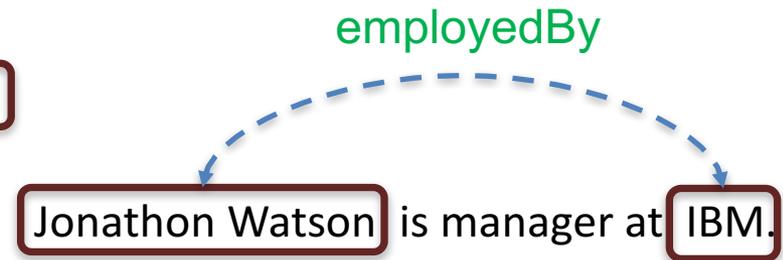
- Supervised methods

Predefined schema (employedBy, bornOn, BirthPlace ...)

Training data



Test data



# Relation extraction

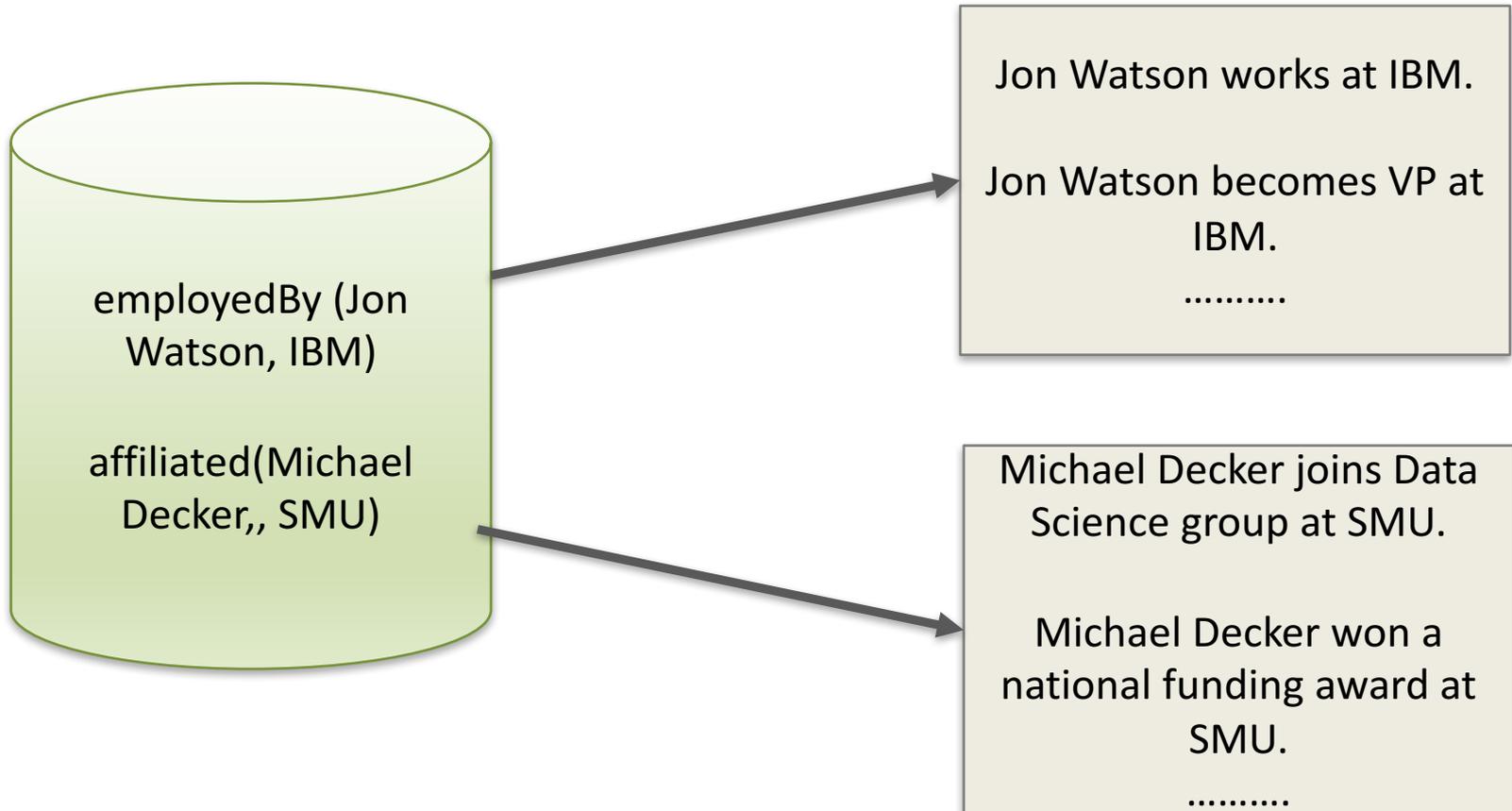
- Supervised methods

Pros: High accuracy and less noise

Cons: Hard and expensive to build labeled data

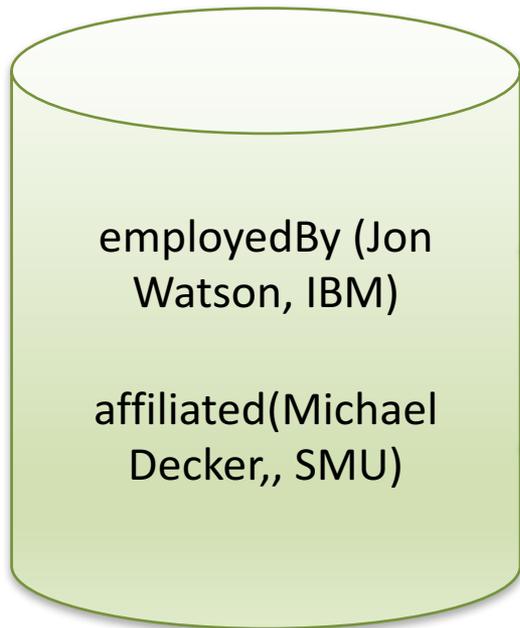
# Relation extraction

- Supervised methods
- Distantly supervised methods



# Relation extraction

- Supervised methods
- Distantly supervised methods



## Training sentences

Jon Watson works at IBM.

Jon Watson becomes VP at  
IBM.

.....

Michael Decker joins Data  
Science group at SMU.

Michael Decker won a  
national funding award at  
SMU.

.....

# Relation extraction

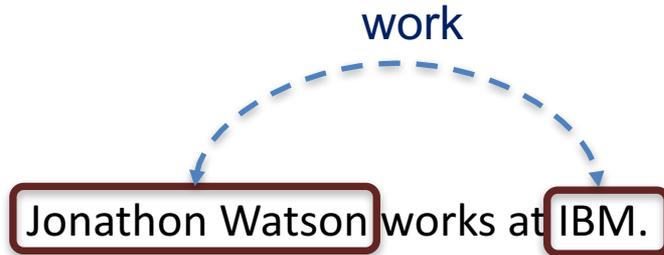
- Supervised methods
- Distantly supervised methods

Pros: Overcome the effort of labeling data

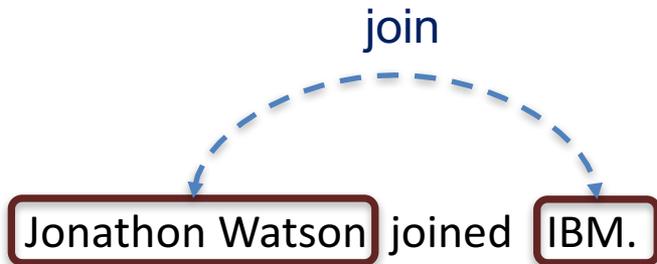
Cons: Dependency of existing knowledge graph and corresponding text

# Relation extraction

- Supervised methods
- Distantly supervised methods
- Unsupervised methods (OpenIE, Universal Schema)



(ROOT (S (NP (Jon Watson)) (VP (VBZ **works**) (PP (IN at) (NP IBM))))



(ROOT (S (NP (Jon Watson)) (VP (VBD **joined**) (NP IBM)))

# Relation extraction

- Supervised methods
- Distantly supervised methods
- Unsupervised methods (OpenIE, Universal Schema)

Pros: eliminates the effort of labeling data

Cons: Noisy, large number of relations

# Relation extraction

- Supervised methods
- Distantly supervised methods
- Unsupervised methods (OpenIE, Universal Schema)

Relation1



Relation2



Relation3



# Relation extraction (Universal Schema)

	<b>employeBy</b>	<b>affiliated</b>	<b>Leader of</b>
Jon	x		x
Michael		x	
Steve		x	x
Joyce	x	x	x

- Clustering using vector similarity
- Matrix completion and fill the empty values [Yao L. et al., 2012]

# Entity types identification (Universal Schema)

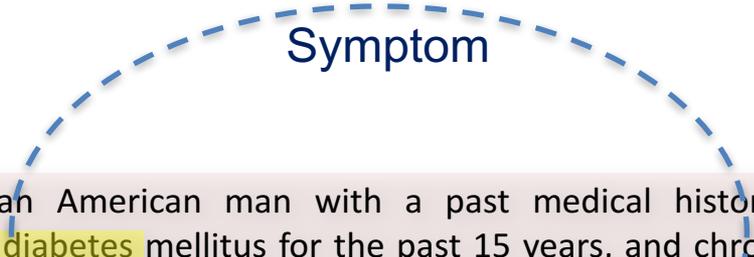
	<b>director</b>	<b>musician</b>	<b>actor</b>
Jon		x	x
Michael	x	x	
Steve			x
Joyce	x		x

- Clustering using vector similarity
- Matrix completion and fill the empty values [Yao L. et al., 2012]

# Relation extraction in domain

- Supervised methods – Need domain experts to label the data
- Distantly supervised methods – Hard to find corresponding text
- Unsupervised methods (OpenIE, Universal Schema) – Noisy

Symptom

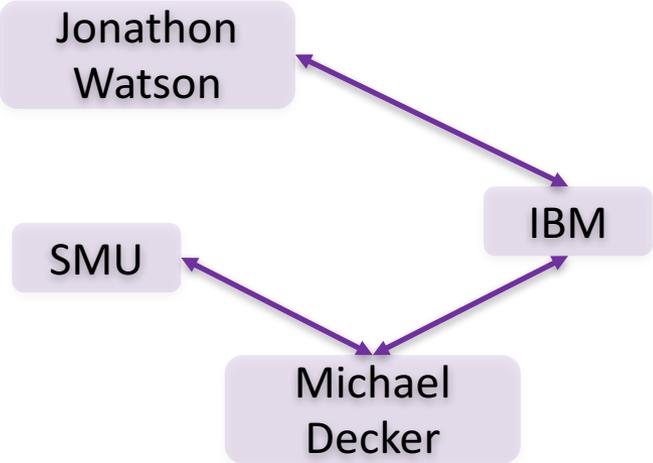
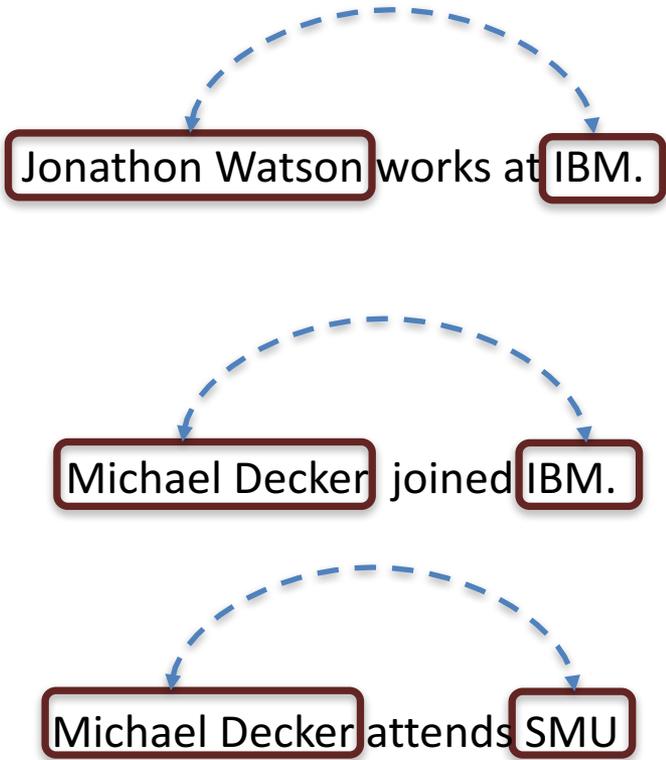


A 59-year-old African American man with a past medical history of hypertension, benign prostatic hypertrophy, type II diabetes mellitus for the past 15 years, and chronic back pain presents to the hospital with gross hematuria. The patient states that he noticed blood in his urine last night. The patient also reports mild, intermittent flank pain. The patient states that his diabetes and blood pressure are well controlled with medications, and that he has managed his chronic back pain with 2 aspirin per day for the past 4 years. Vital signs are Temp- 98.6°F, BP- 124/82 mm/Hg, pulse- 88/min, and RR- 14/min. Blood work is notable for HbA1C of 6.5%. A pyelogram reveals a ring sign. His current fasting glucose is 140mmol/L. What is the most likely etiology of hematuria in this patient?

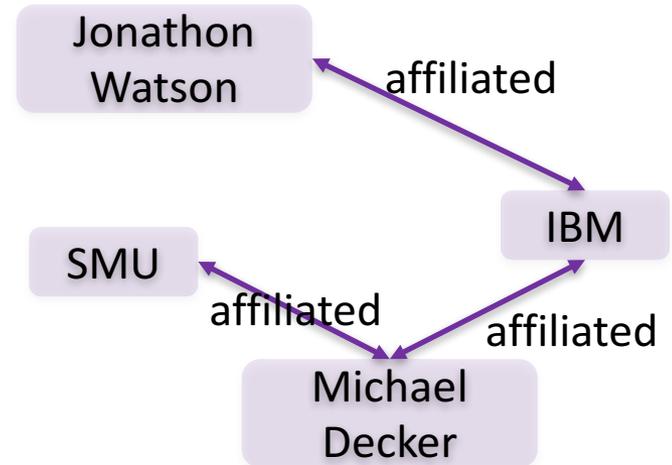
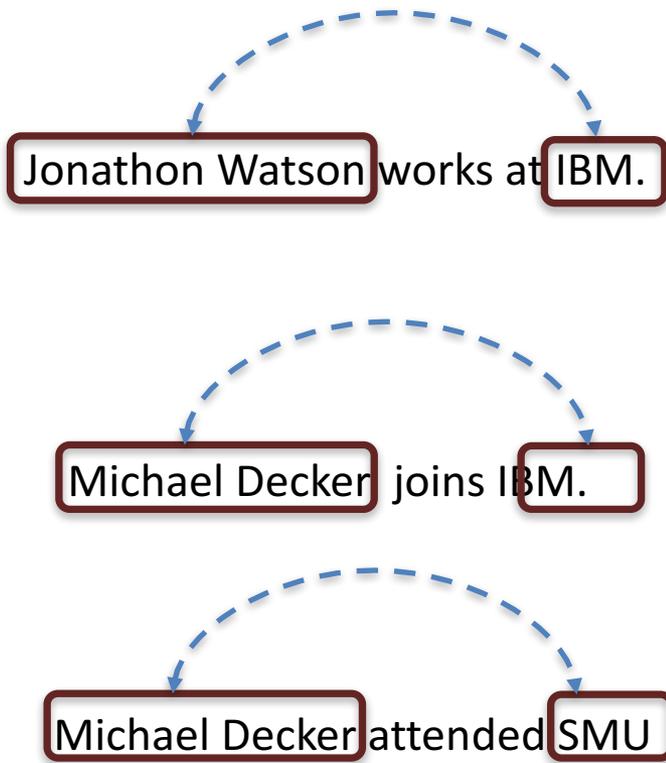
# Knowledge graphs in domain

- Domain specific entity extraction is more challenging
- Limited relation types
- Less explicit mention of entity and relation types in text
- Creating simple schema requires domain experts

# Knowledge graph - Simple



# Knowledge graph – Simple + Schema

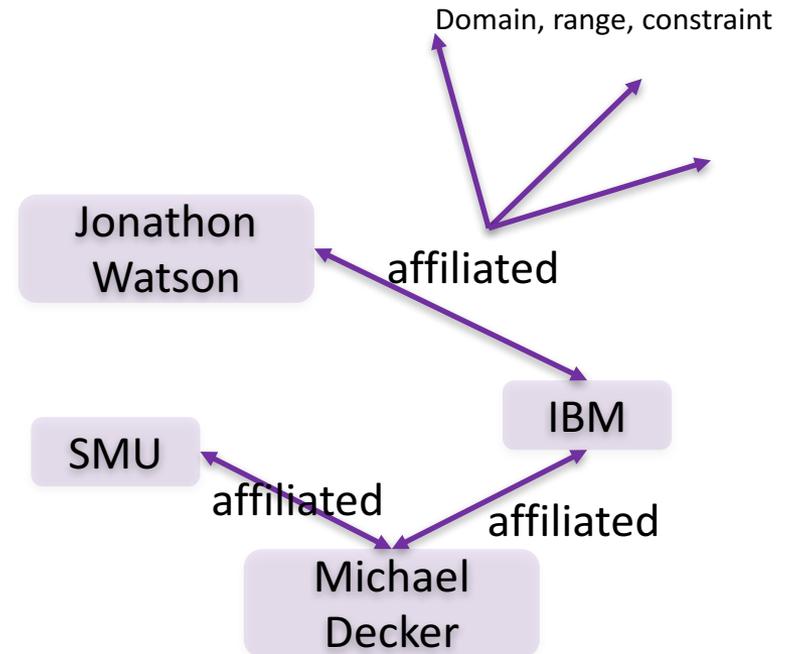


# Knowledge graph – Simple + Schema + Ontology

Jonathon Watson works at IBM.

Michael Decker joined IBM.

Michael Decker attends SMU



# Summary

- Simple knowledge graph works for many applications
- Identify the requirement before finding the solution.
- Many knowledge graphs are publically available

<https://www.youtube.com/watch?v=kao05Arliok&feature=youtu.be>