



EAS

ENVI ANALYTICS SYMPOSIUM

Data Acquisition through Geospatial Information Retrieval

John Farrell
L3Harris Technologies

Principle Investigator, Geospatial AI Analytics IRAD

john.farrell@l3harris.com

Introduction



Wouldn't it be nice if you could say “Alexa, find labels and annotations at 39.7499° N, 104.9917° W, and show me the results on my map”. This may be possible one day using Geospatial Information Retrieval or GIR.

Introduction



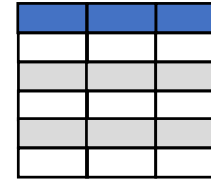
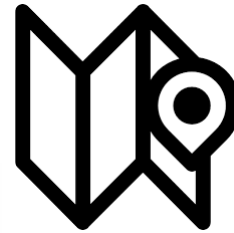
Wouldn't it be nice if you could say “Alexa, find labels and annotations at 39.7499° N, 104.9917° W, and show me the results on my map”. This may be possible one day using Geospatial Information Retrieval or GIR.

- What can I learn about a location from the Internet?
- How can I augment my geospatial understanding of a location?
- How can I do this automatically?

A note about structure and search



- Structured, Semi-structured and Unstructured data
 - Structured
 - Semi-Structured
 - Unstructured
- Search and Query
 - Query is precise, complete
 - Search is imprecise, ranked

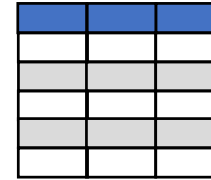
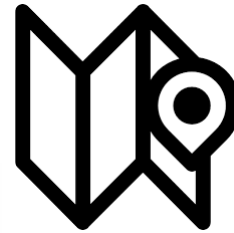


OGC[®]
Making location count.
www.opengeospatial.org

A note about structure and search



- Structured, Semi-structured and Unstructured data
 - Structured
 - Semi-Structured
 - Unstructured
- Search and Query
 - Query is precise, complete
 - Search is imprecise, ranked



OGC[®]
Making location count.
www.opengeospatial.org

- Much of the information on the Internet contains references to places on earth.
- Traditionally Geographic Information Systems (GIS) focus on geospatial information available as structured data (OGC, KML, GML, etc.).
- Growing amounts of spatial references are available in unstructured or semi-structured data available today.

For Example



Just the first paragraph contains at least 16 phrases with a reference to a location

Denver

From Wikipedia, the free encyclopedia

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

"Mile-High City" redirects here. For the capital of Lesotho, see [Maseru](#).

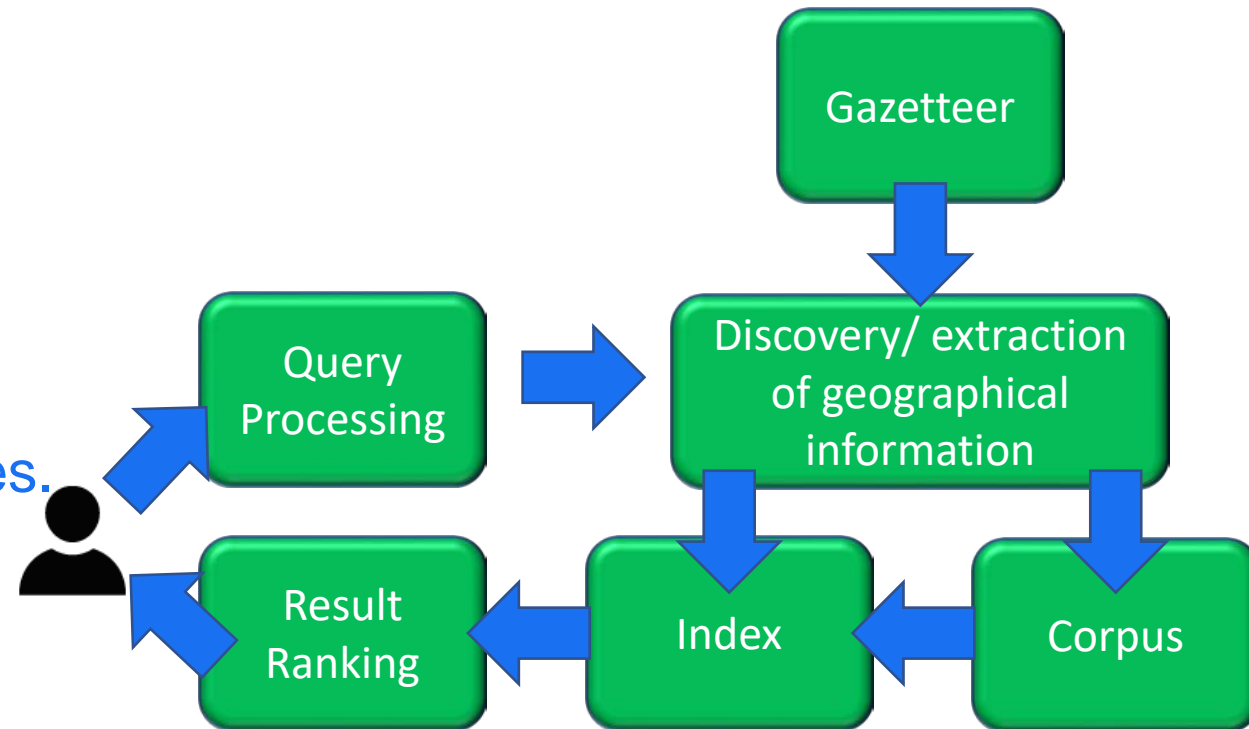
Denver (/ˈdɛnvər/), officially the **City and County of Denver**, is the [capital](#) and [most populous municipality](#) of the U.S. state of [Colorado](#). Denver is located in the [South Platte River Valley](#) on the western edge of the [High Plains](#) just east of the [Front Range](#) of the [Rocky Mountains](#). The Denver downtown district is immediately east of the confluence of [Cherry Creek](#) with the [South Platte River](#), approximately 12 mi (19 km) east of the [foothills](#) of the [Rocky Mountains](#). [Denver](#) is named after [James W. Denver](#), a governor of the [Kansas Territory](#), and it is [nicknamed](#) the [Mile High City](#) because its official elevation is exactly one mile (5280 feet or 1609.3 meters) above sea level.^[17] The [105th meridian west of Greenwich](#), the longitudinal reference for the [Mountain Time Zone](#), passes directly through [Denver Union Station](#).

Denver is ranked as a [Beta world city](#) by the [Globalization and World Cities Research Network](#). With an estimated population of 716,492 in 2018, Denver is the [19th-most populous U.S. city](#), and with a 19.38% increase since the [2010 United States Census](#), it has been one of the fastest-growing major cities in the United States.^[18] The 10-county [Denver-Aurora-Lakewood, CO Metropolitan Statistical Area](#) had an estimated 2018 population of 2,932,415 and is the [19th most populous U.S. metropolitan statistical area](#).^[19] The 12-city [Denver-Aurora, CO Combined Statistical Area](#) had an estimated 2018 population of 3,572,798 and is the [15th most populous U.S. metropolitan area](#).^[20] Denver is the most populous city of the 18-county [Front Range Urban Corridor](#), an oblong urban region stretching across two states with an estimated 2018 population of 4,976,781.^[21] Denver is the most populous city within a 500-mile (800 km) radius and the second-most populous city in the [Mountain West](#) after [Phoenix, Arizona](#). In 2016, Denver was named the best place to live in the United States by [U.S. News & World Report](#).^[22]

Basic approach



- Query Processing: build a query that represents location and space using natural language.
- Discovery/extraction of geographical information: find and formulate structured geographical data from the text.
- Gazetteers: catalogs of place name related information, including geographic coordinates.
- Corpus: source of our content.
- Index: the 'card catalog' for the extracted geographic data.
- Result Ranking: Results ranked by some priority determination.



We already use GIR



- Search engines have used local search in recent years.
- Works well with commercial directory listings and travel pages.
- Google has program to get your business listed and located.



We already use GIR



- Search engines have used local search in recent years.
- Works well with commercial directory listings and travel pages.
- Google has program to get your business listed and located.

What if I want something not in a business directory, or travel ad?

- The web provides a very large amount unstructured web content that refers to geographical information in semi-structured text.
- Easy to retrieve if there is a direct match between our query and the terms used in the web page.
- This presentation will **introduce** approaches to developing methods to find, index, rank and retrieve geographical data from text.

Understanding geographical information in text



- Consider the query “attractions near the Ritz-Carlton Denver”. It consists of three important parts
 - A theme (attractions)
 - A spatial relationship (near)
 - A location (Ritz-Carlton Denver).

Understanding geographical information in text



- Consider the query “attractions near the Ritz-Carlton Denver”. It consists of three important parts
 - A theme (attractions)
 - A spatial relationship (near)
 - A location (Ritz-Carlton Denver).
- It is likely that by Denver we mean Denver, CO, but this is an assumption.

Understanding geographical information in text



- Consider the query “attractions near the Ritz-Carlton Denver”. It consists of three important parts
 - A theme (attractions)
 - A spatial relationship (near)
 - A location (Ritz-Carlton Denver).
- It is likely that by Denver we mean Denver, CO, but this is an assumption.
- What do we mean by near? Next door? Walking distance? Driving distance?

Understanding geographical information in text



- Consider the query “attractions near the Ritz-Carlton Denver”. It consists of three important parts
 - A theme (attractions)
 - A spatial relationship (near)
 - A location (Ritz-Carlton Denver).
- It is likely that by Denver we mean Denver, CO, but this is an assumption.
- What do we mean by near? Next door? Walking distance? Driving distance?
- Are attractions ranked how close they are to the centroid of the Denver city limits? The centroid of ‘downtown’ Denver? The ratings of the attractions?

Understanding geographical information in text



- Consider the query “attractions near the Ritz-Carlton Denver”. It consists of three important parts
 - A theme (attractions)
 - A spatial relationship (near)
 - A location (Ritz-Carlton Denver).
- It is likely that by Denver we mean Denver, CO, but this is an assumption.
- What do we mean by near? Next door? Walking distance? Driving distance?
- Are attractions ranked how close they are to the centroid of the Denver city limits? The centroid of ‘downtown’ Denver? The ratings of the attractions?
- What are the attractions? A monument (point), state park (polygon) or a scenic drive (polyline)?

Understanding geographical information in text



- Consider the query “attractions near the Ritz-Carlton Denver”. It consists of three important parts
 - A theme (attractions)
 - A spatial relationship (near)
 - A location (Ritz-Carlton Denver).
- It is likely that by Denver we mean Denver, CO, but this is an assumption.
- What do we mean by near? Next door? Walking distance? Driving distance?
- Are attractions ranked how close they are to the centroid of the Denver city limits? The centroid of ‘downtown’ Denver? The ratings of the attractions?
- What are the attractions? A monument (point), state park (polygon) or a scenic drive (polyline)?
- Theme/Spatial Relationship/Location is a rule. More on that later.

Geographic references



We need to find geographic references in text, and unambiguously resolve these to locations on the earth's surface.

Ritz-Carlton Denver

From Wikipedia, the free encyclopedia

The Ritz-Carlton Denver, formerly known as the **Embassy Suites Downtown**, is a [skyscraper](#) in [Denver, Colorado](#). The building, a part of the [Denver Place](#) complex, was completed in 1983, and rises 38 floors and 390 feet (119 m) in height.^[1] The building stands as the eighteenth-tallest building in Denver and Colorado.

The building was constructed as **Embassy Suites and Apartments at Denver Place** in 1983,^[2] at which point it was the second tallest hotel building in Denver, behind the 522-foot (159 m) [MCI Building](#), now known as [707 17th Street](#); this structure is a mixed use hotel and commercial property whose lowest 20 floors are occupied by a Private Condominiums in the tower, including several [suites](#).^[1] The uppermost 17 floors of the tower consist of [apartments](#). In 2005, the building was converted into a wholly owned condominiums (floors 15-19).^[1] This conversion was completed in early 2008, and the hotel had its grand opening

General information	
Status	Complete
Type	Hotel, Residential ^[1]
Location	1881 Curtis Street, Denver, Colorado, United States
Coordinates	 39°45′01″N 104°59′31″W
Opening	1983 ^[2]
Owner	Amerimar Enterprises, Inc. ^[1]
Height	
Roof	390 ft (120 m) ^[1]
Technical details	
Floor count	38 ^[2]
Design and construction	
Developer	Amerimar Enterprises, Inc. , CJS Denver Holdings, LLC ^[1]

See also [edit]

- List of tallest buildings in Denver

References [edit]

- ↑ *abcdefg* "[The Ritz-Carlton, Denver](#)". Emporis.com. Retrieved 2008-06-16.
- ↑ *abc* "[Embassy Suites Downtown](#)". SkyscraperPage.com. Retrieved 2008-06-16.
- ↑ "[707 17th Street](#)". Emporis.com. Retrieved 2008-06-16.
- ↑ "[The tallest buildings in Colorado](#)". DenverSkyscrapers.com. Archived from the original on 2008-03-28. Retrieved 2008-06-15.

External links [edit]

- [Official website for The Ritz-Carlton, Denver](#)

Step 1: Identify geographical references unambiguously



The located object and the reference object can be given place names that can vary in geographical granularity

GeoNames licensed under a Creative Commons Attribution 4.0 License

Research

Research Home > Tools > Getty Thesaurus of Geographic Names > Search Results

Getty Thesaurus of Geographic Names® On Search Results

[New Search](#) [Previous Page](#)

Find Name: **denver**

Place Type:

Nation:

[View Selected Records](#) [Select All Records](#) [Clear All](#)

Click the icon to view the hierarchy. Check boxes to view multiple records at once.

- ☐ **Big Denver Creek** (creek)
(World, North and Central America, United States, Alaska, Yukon-K...
- ☐ **Curlee** (inhabited place)
(World, North and Central America, United States, Tennessee, Can...
- ☐ **Denver** (inhabited place)
(World, Europe, United Kingdom, England, Norfolk county) [7028355]
- ☐ **Denver** (inhabited place)
(World, North and Central America, United States, Alabama, Tallapoosa county) [2284244]
- ☐ **Denver** (inhabited place)
(World, North and Central America, United States, Arkansas, Carroll county) [2008098]
- ☐ **Denver** (county)
(World, North and Central America, United States, Colorado) [2000207]
Denver county

denver all countries

search [advanced search](#)

2053 records found for "denver"

Name	Country	Feature class	Latitude	Longitude
1 Denver DEN,Dehnrer,Denber,Denr,r,Denver,Denver County,Denvera,Denveres,Denveris,Denverium,Denvero,Denverès,...	United States , Colorado Denver	seat of a first-order administrative division population 682,543, elevation 1609m	N 39° 44' 20"	W 104° 59' 4"
2 Denver International Airport Aeroport Internacional de Denver,Aeroport international de Denver,Aeroporto Internacional de Denver,...	United States , Colorado Denver	airport elevation 1637m	N 39° 51' 32"	W 104° 40' 23"
3 Denver Comte de Denver,Comité de Denver,Denver	United States Denver			
4 Denver	Australia Hepburn			
5 Denver Athletic Club	United States Denver			
6 City Park City Park,Denver City Park	United States Denver			
7 Sloan Lake Sloan Lake,Sloan's Lake,Sloans Lake	United States Denver			
8 Washington Park	United States Denver			
9 Cheesman Park	United States Denver			
10 Cranmer Park	United States Denver			
11 Civic Center Park	United States Denver			

WIKIPEDIA
The Free Encyclopedia

[Main page](#)
[Contents](#)
[Featured content](#)
[Current events](#)
[Random article](#)
[Donate to Wikipedia](#)
[Wikipedia store](#)

[Interaction](#)
[Help](#)
[About Wikipedia](#)
[Community portal](#)
[Recent changes](#)
[Contact page](#)

[What's here](#)
[Related changes](#)

Article [Talk](#)

Ritz-Carlton Denver

From Wikipedia, the free encyclopedia

Article [Talk](#)

Denver (disambiguation)

From Wikipedia, the free encyclopedia

Denver is the capital of the U.S. state of Colorado.

Denver may also refer to:

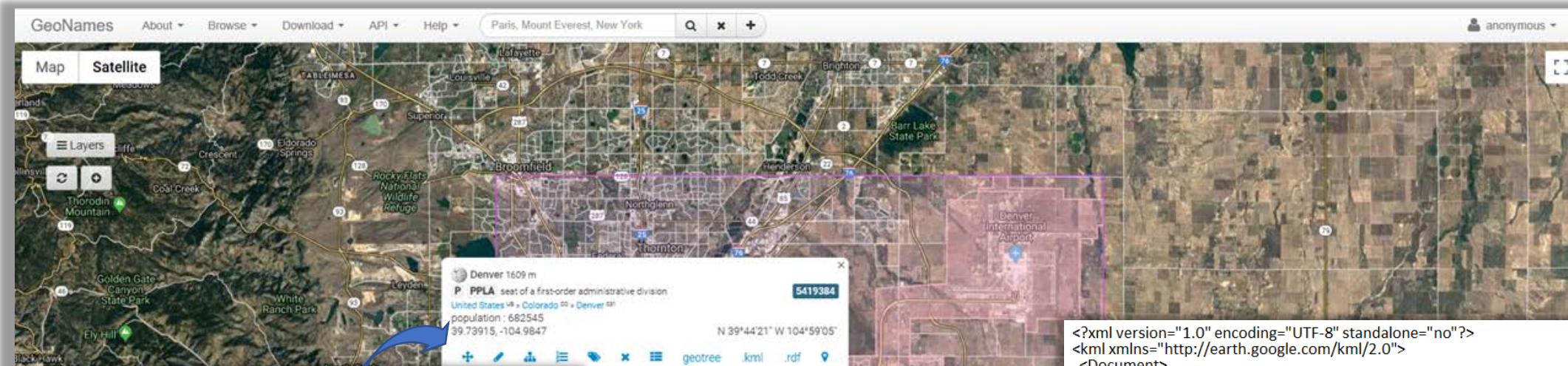
Places [\[edit \]](#)

- [Denver, Victoria](#), a locality in Australia
- [Denver, Nova Scotia](#), a village in Canada
- [Denver, Norfolk](#), a village and civil parish in the United Kingdom

United States [\[edit \]](#)

- [Denver Basin](#), a geologic basin in eastern Colorado
- [Denver, Illinois](#), an unincorporated community
- [Denver Township, Richland County, Illinois](#)
- [Denver, Indiana](#), a town
- [Denver, Iowa](#), a city
- [Denver, Kentucky](#), an unincorporated community

GeoNames



5419384 Denver

Feature Hierarchy History Tags Alternate names

Administrative hierarchy Parents Children

country	United States (US)	▼	↕
adm1	Colorado (CO)	▼	↕
adm2	Denver (031)	▼	↕
adm3	no admin3	▼	↕
adm4	no admin4	▼	↕
adm5	no admin5	▼	↕

✓ ✕

Containment

```
<?xml version="1.0" encoding="UTF-8" standalone="no"?>
<kml xmlns="http://earth.google.com/kml/2.0">
  <Document>
    <Folder>
      <description>more than 8 million features</description>
      <name>geonames.org</name>
      <visibility>1</visibility>
      <open>1</open>
      <LookAt>
        <longitude>-104.9847</longitude>
        <latitude>39.73915</latitude>
        <range>44499.31740609377</range>
      </LookAt>
      <NetworkLink>
        <description/>
        <name>loading data from geonames.org</name>
        <visibility>1</visibility>
        <open>1</open>
        <refreshVisibility>1</refreshVisibility>
        <flyToView>0</flyToView>
        <Url>
          <href>http://www.geonames.org/servlet/geonames?srv=2&type=kml&numrows=30...</href>
          <refreshInterval>2</refreshInterval>
          <viewRefreshMode>onStop</viewRefreshMode>
          <viewRefreshTime>1</viewRefreshTime>
        </Url>
      </NetworkLink>
    </Folder>
  </Document>
</kml>
```


Place names and spatial references are ambiguous



mile high city

search

[advanced search](#)

all countries

6 records found for "mile high city"

Name	Country	Feature class	Latitude	Longitude
<div>1</div> <div><div><div></div><div>Denver</div><div></div></div><div>DEN,Dehnver,Denber,Denv"r,Denver,Denver County,Denvera,Denveres,Denveris,Denverium,Denvero,Denverés,...</div></div>	<div>United States, Colorado</div> <div>Denver</div>	<div>seat of a first-order administrative division</div> <div>population 682,543, elevation 1609m</div>	<div>N 39° 44' 20"</div>	<div>W 104° 59' 4"</div>
<div>2</div> <div><div><div></div><div>Lead</div><div></div></div><div>Lead,Lead City,Lid,Mile High City,Washington,lai de,lyd,lyd dakortay jnwby,لید،لید</div></div>				

Top things to do in Denver



The Green Dragon
Shopping



Bucher's Mill
Covered Bridge
Historical landmark



Reinholds Station
Trinity Chapel
Historical landmark



Reamstown
Memorial Park
WWII memorial &
sports facilities

[Denver travel guide](#)

Denver, Pennsylvania - Wikipedia

https://en.wikipedia.org/wiki/Denver,_Pennsylvania

Denver is a borough in Lancaster County, Pennsylvania, United States. The population was 3,332 at the 2000 census. Contents. 1 History; 2 Geography ...

County: Lancaster Elevation: 394 ft (120 m)
ZIP code: 17517 Area code(s): 717 Exchanges: 335, 336

[History](#) · [Geography](#) · [Demographics](#) · [Education](#)

Denver, PA - Denver, Pennsylvania Map & Directions - MapQuest

<https://www.mapquest.com/us/pennsylvania/denver-pa-282022119>

Get directions, maps, and traffic for Denver, PA. Check flight prices and hotel availability for your visit.



Denver

Denver is a borough in Lancaster County, Pennsylvania, United States. The population was 3,332 at the 2000 census. [Wikipedia](#)

Weather: 84°F (29°C), Wind S at 2 mph (3 km/h), 77% Humidity

Zip code: 17517

Hotels: 3-star averaging \$145. [View hotels](#)

Getting there: 1 h 19 min flight, from \$333. [View flights](#)

Population: 3,879 (2017)

Events

Sun, Jul 21
9:00 AM
Motorcycle Benefit Ride For Payton Zimmer...
Fivepointville Fire Company

Fri, Jul 19
7:30 AM
Wwe Reading Tickets
Santander Arena

In the neighborhood
Next to
Foothills of
Near
Down the road
Downtown
Located in
Western edge of
East of
Passes through

Automating geographical reference identification



- There are three basic approaches to identifying candidate georeferences:
 - List lookup
 - Need exact match
 - Can use capitalization, stemming, thesaurus, etc.
 - Manual

Buildings -> buildings -> building



Automating geographical reference identification



- There are three basic approaches to identifying candidate georeferences:

- List lookup

- Need exact match
- Can use capitalization, stemming, thesaurus, etc.
- Manual

Buildings -> buildings -> building



- Knowledge or rule-based methods

- Take context into account
- Natural Language Processing (NLP)
- More general than list lookup
- Tokenization, stemming, stop words, named entity resolution (starts to use machine learning)

Geographic statement = Theme/Spatial Relationship/Location

Automating geographical reference identification



- There are three basic approaches to identifying candidate georeferences:

- List lookup

- Need exact match
 - Can use capitalization, stemming, thesaurus, etc.
 - Manual

Buildings -> buildings -> building



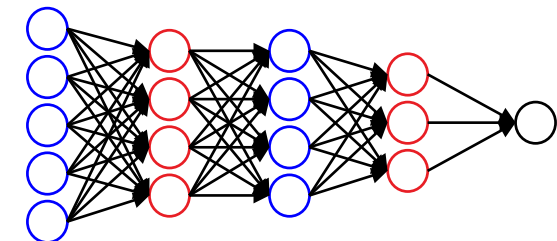
- Knowledge or rule-based methods

- Take context into account
 - Natural Language Processing (NLP)
 - More general than list lookup
 - Tokenization, stemming, stop words, named entity resolution (starts to use machine learning)

Geographic statement = Theme + Spatial Relationship + Location

- Machine learning

- Supervised and unsupervised
 - Less manual effort (with exception of labeling data)
 - Topic modeling, deep encoding, segmentation, classification



- 



Conclusion



- Geospatial Information Retrieval (GIR) is still in the early stages, but promising
- Hopefully you gained some understanding of Geospatial Information Retrieval
 - Geospatial Information Retrieval is fairly easy for humans
 - Automating Geospatial Information Retrieval is hard, but there has been progress

Questions?