

Object extraction techniques and visual image search with Semantic web techniques

Aninda Maulik

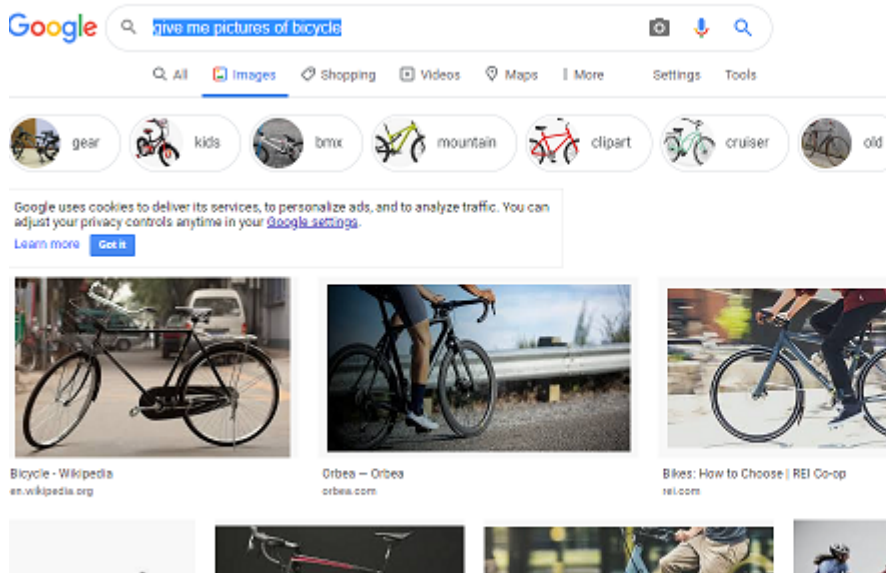
Supervisors: Prof. Pierre Maret
Dennis Diefenbach

Cyber Physical and Social Systems
University of Jean Monnet

July 2020

Introduction: Just Google

- images with bicycles



images with bicycles on the right



images with bicycles on the right



All

Images

News

Videos

Maps

More

Settings

Tools

Collections

road

size

riding

cycling

bike shops

rob english

saddle

road bike fit

mountain bike

lock

bike locks

frame

wh



How to pick the right size ...
pinterest.com



Bikes v. Cars: Who Has the Right of Way ...
lorenzoflorenzoni.com



Right Bike measures you up for the ...
newatlas.com



Road Bike Frame Sizes ...
pinterest.com



How to Pick the Best B...
dailyburn.com



6 Relatively Cheap Bikes You Can...
forbes.com

- QAnswer is knowledge based QA system.
- A knowledge base is a collection of facts that can be interpreted by a machine
- Such a fact can look like this:"bicycle" "in" "right".
- QAnswer translates from a natural language question to correct SPARQL queries.

- Implementation of an Algorithm for object extraction.
- Design of a semantic web modelling for extracted data.
- Implementation of a visual image search engine through Qanswer.

Implementation of an Algorithm for object extraction.

YOLO-(You Only Look Once)

- Implementation of an Algorithm for object extraction.



Computer
Vision API
with
PyTorch
Framework
and use
YOLO



| | | | | | |
|----|----|-----|-----|----|-------------------------|
| 84 | 34 | 289 | 170 | 20 | 0.934078 |
| 4 | 36 | 88 | 162 | 20 | 0.841598 |
| X1 | Y1 | X2 | Y2 | | |
| | | | | | class number/class name |
| | | | | | confidence |

Class Number, Class Name, QID

| | | |
|--------------------------|---------------------------|--------------------------|
| 0. person-Q215627 | 25.umbrella-Q41607 | 49.orange-Q39338 |
| 1.bicycle-Q11442 | 26.handbag-Q467505 | 50.broccoli-Q47722 |
| 2.car-Q1420 | 27.tie-Q44416 | 51.carrot-Q81 |
| 3.motorbike-Q34493 | 28.suitcase-Q200814 | 52.hot dog-Q181055 |
| 4.aeroplane-Q197 | 29.frisbee-Q131689 | 53.pizza-Q177 |
| 5.bus-Q5638 | 30.skis-Q172226 | 54.donut-Q192783 |
| 6.train-Q870 | 31.snowboard-Q178131 | 55.cake-Q13276 |
| 7.truck-Q43193 | 32.sports ball-Q63347096 | 56.chair-Q15026 |
| 8.boat-Q35872 | 33.kite-Q107061 | 57.sofa-Q131514 |
| 9.traffic light-Q8004 | 34.baseball bat-Q809910 | 58.pottedplant-Q203834 |
| 10.fire hydrant-Q634299 | 35.baseball glove-Q809894 | 59.bed-Q42177 |
| 11.stop sign-Q250429 | 36.skateboard-Q15783 | 60.diningtable-Q10578291 |
| 12.parking meter-Q953960 | 37.surfboard-Q457689 | 61.toilet-Q7857 |
| 13.bench-Q204776 | 38.tennis racket-Q153362 | 62.tvmonitor-Q289 |
| 14.bird-Q5113 | 39.bottle-Q80228 | 63.laptop-Q3962 |
| 15.cat-Q4167836 | 40.wine glass-Q1531435 | 64.mouse-Q7987 |
| 16.dog-Q144 | 41.cup-Q81727 | 65.remote-Q185091 |
| 17.horse-Q726 | 42.fork-Q81881 | 66.keyboard-Q250 |
| 18.sheep-Q7368 | 43.knife-Q32489 | 67.cell phone-Q17517 |
| 19.cow-Q830 | 44.spoon-Q81895 | 68.microwave-Q127956 |
| 20.elephant-Q7378 | 45.bowl-Q153988 | 69.oven-Q36539 |
| 21.bear-Q30090244 | 46.banana-Q503 | 70.toaster-Q14890 |
| 22.zebra-Q32789 | 47.apple-Q89 | 71.sink-Q140565 |
| 23.giraffe-Q862089 | 48.sandwich-Q28803 | 72.refrigerator-Q37828 |
| 24.backpack-Q5843 | | |

Class Number, Class Name, QID

```
https://github.com/pjreddie/darknet/blob/master/data/  
coco.names
```

```
73.book-Q571
```

```
74.clock-Q376
```

```
75.vase-Q191851
```

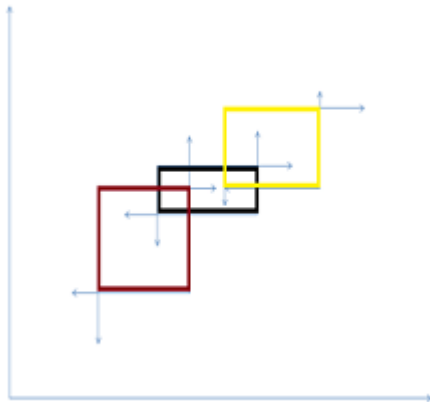
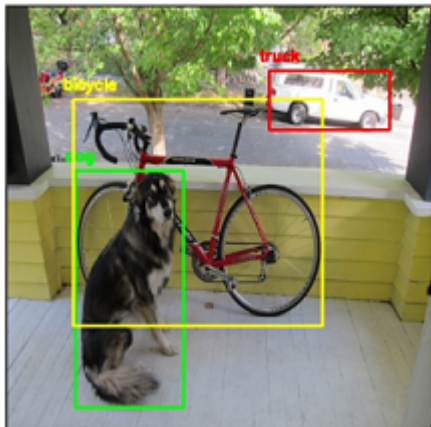
```
76.scissors-Q40847
```

```
77.teddy bear-Q213477
```

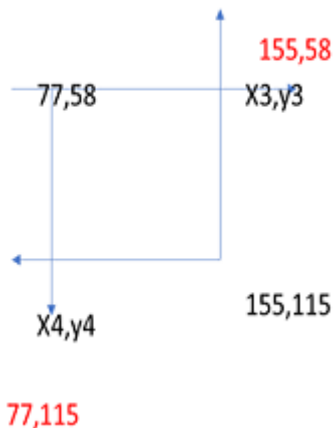
```
78.hair drier-Q15004
```

```
79.toothbrush-Q134205
```


Bounding Box



Co-ordinate representation of Bounding Box



$x_3 = x_2 = 155, y_3 = y_1 = 58;$
 $x_4 = x_1 = 77, y_4 = y_2 = 115$

Image-Object Relations

| Image | relation | property value |
|-------|-------------------|----------------|
| | has on the left | |
| | has on the right | |
| | has on the top | |
| | has on the bottom | |
| | has in the center | |

Algorithms for Image-Object Relations

Algorithm 1 has on the left and right

```
1: if  $X - \text{centre} \leq 0.3 * X - \text{ImageDimensions}$  then  
2:    $\text{hasontheleft} \leftarrow \text{object}$   
3: else  
4:   if  $X - \text{centre} \geq 0.6 * X - \text{ImageDimensions}$  then  
5:      $\text{hasontheright} \leftarrow \text{object}$   
6:   end if  
7: end if
```

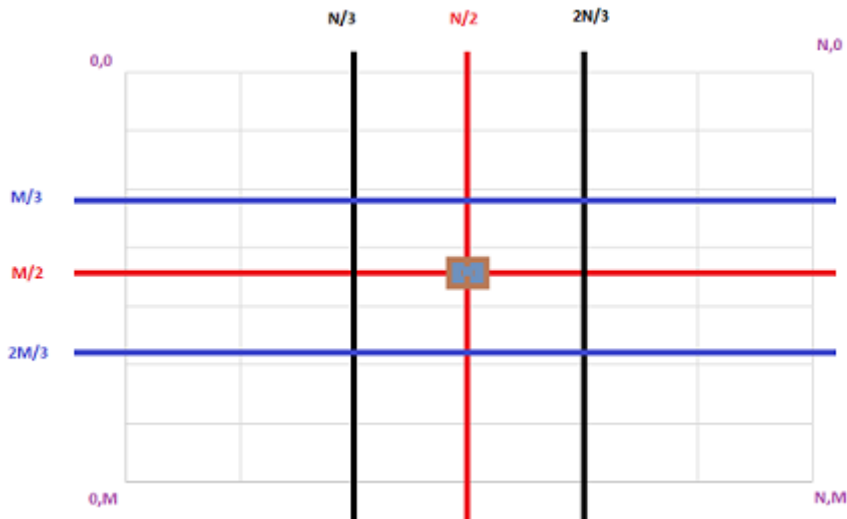
Algorithm 2 has on the top and bottom

```
1: if  $Y - \text{centre} \leq 0.3 * Y - \text{ImageDimensions}$  then  
2:    $\text{hasonthetop} \leftarrow \text{object}$   
3: else  
4:   if  $Y - \text{centre} \geq 0.6 * Y - \text{ImageDimensions}$  then  
5:      $\text{hasonthebottom} \leftarrow \text{object}$   
6:   end if  
7: end if
```

Algorithm 3 has in the center

```
1: if  $X - \text{centre} \geq 0.3 * X - \text{ImageDimensions},$   
    $X - \text{centre} \leq 0.66 * X - \text{ImageDimensions},$   
    $Y - \text{centre} \geq 0.3 * Y - \text{ImageDimensions},$   
    $Y - \text{centre} \leq 0.66 * Y - \text{ImageDimensions}$  then  
2:    $\text{hasinthecenter} \leftarrow \text{object}$ 
```

An attempt to show a graphical representation of top, bottom, left, right, and central position



Algorithms for Image-Object Relations

Algorithm 1 has on the left and right

```
1: if  $X - \text{centre} \leq 0.3 * X - \text{ImageDimensions}$  then  
2:    $\text{hasontheleft} \leftarrow \text{object}$   
3: else  
4:   if  $X - \text{centre} \geq 0.6 * X - \text{ImageDimensions}$  then  
5:      $\text{hasontheright} \leftarrow \text{object}$   
6:   end if  
7: end if
```

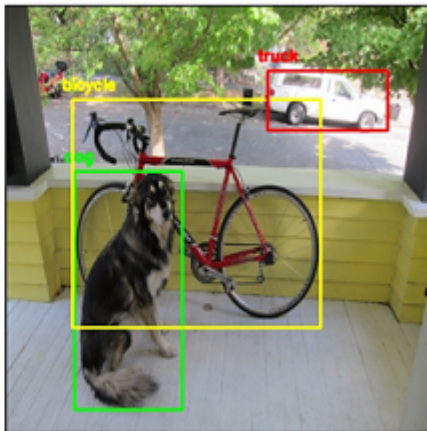
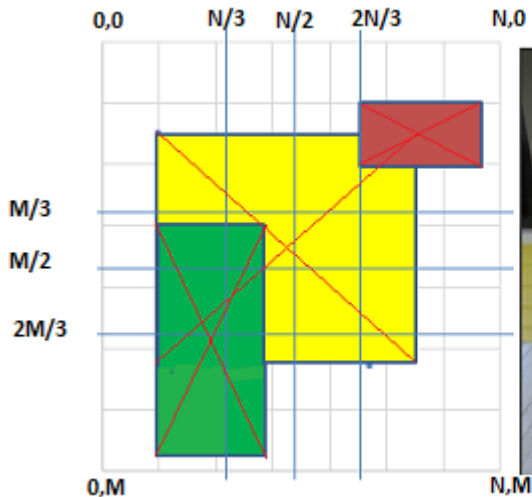
Algorithm 2 has on the top and bottom

```
1: if  $Y - \text{centre} \leq 0.3 * Y - \text{ImageDimensions}$  then  
2:    $\text{hasonthetop} \leftarrow \text{object}$   
3: else  
4:   if  $Y - \text{centre} \geq 0.6 * Y - \text{ImageDimensions}$  then  
5:      $\text{hasonthebottom} \leftarrow \text{object}$   
6:   end if  
7: end if
```

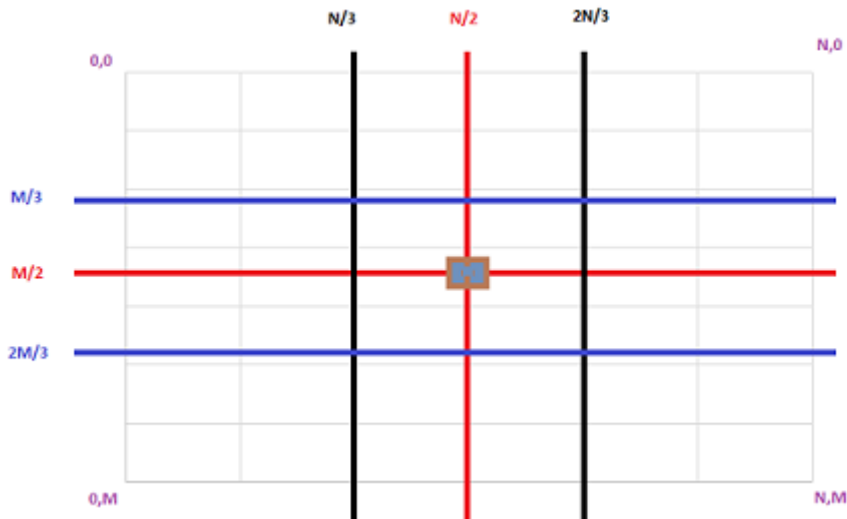
Algorithm 3 has in the center

```
1: if  $X - \text{centre} \geq 0.3 * X - \text{ImageDimensions},$   
    $X - \text{centre} \leq 0.66 * X - \text{ImageDimensions},$   
    $Y - \text{centre} \geq 0.3 * Y - \text{ImageDimensions},$   
    $Y - \text{centre} \leq 0.66 * Y - \text{ImageDimensions}$  then  
2:    $\text{hasinthecenter} \leftarrow \text{object}$ 
```

An attempt to show a graphical representation of Image-Object relation in comparison to an incoming photo



An attempt to show a graphical representation of Image-Object relation



Using Image-Object relation we create a csv file

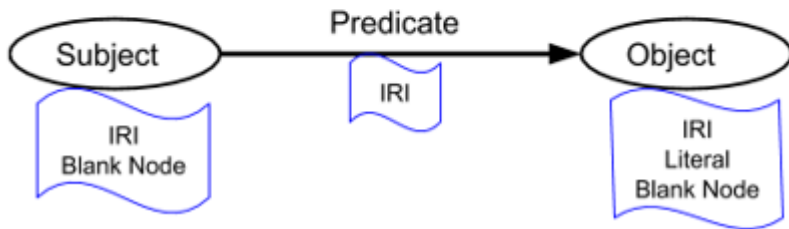
| X1 | Y1 | X2 | Y2 | object name | Image name | has on the left | has on the right | has on the top | has on the bottom | has in the center |
|------|------|------|------|-------------|------------|-----------------|------------------|----------------|-------------------|-------------------|
| 190 | 813 | 3897 | 1932 | airplane | Antonov_20 | na | airplane | na | airplane | airplane |
| 220 | 596 | 5021 | 1673 | airplane | EBACE_20 | na | airplane | na | airplane | airplane |
| 742 | 1303 | 3933 | 2099 | airplane | Embraer_1 | na | airplane | na | airplane | airplane |
| 889 | 1035 | 1378 | 1374 | airplane | Kirchthurm | na | airplane | na | airplane | airplane |
| 172 | 278 | 4018 | 1722 | airplane | Lufthansa | na | airplane | na | airplane | airplane |
| 331 | 532 | 2362 | 1245 | airplane | North_Arr | na | airplane | na | airplane | airplane |
| 1355 | 704 | 1444 | 813 | person | North_Arr | na | person | na | person | person |
| 833 | 251 | 2202 | 1761 | airplane | Paris_Air | na | airplane | na | airplane | airplane |
| 1460 | 1207 | 1579 | 1738 | person | Playing_in | na | person | na | person | na |
| 224 | 1063 | 466 | 1813 | person | Playing_in | person | na | na | person | na |
| 756 | 1286 | 845 | 1562 | person | Playing_in | person | na | na | person | na |
| 2374 | 990 | 2486 | 1359 | person | Playing_in | na | person | na | person | na |
| 468 | 1194 | 546 | 1264 | frisbee | Playing_in | frisbee | na | na | frisbee | na |
| 264 | 273 | 2284 | 1571 | airplane | RUAG_Avina | na | airplane | na | airplane | airplane |

Design of a semantic web modelling for extracted data.

Image-Object

- Design of a semantic web modelling based on the csv file.

| Image | relation | property value |
|-------|-------------------|----------------|
| | has on the left | |
| | has on the right | |
| | has on the top | |
| | has on the bottom | |
| | has in the center | |



Object-Object Relation

| Object | relation | property value |
|--------|------------------|----------------|
| | left of | object |
| | right of | object |
| | top of | object |
| | bottom of | object |
| | close to | object |
| | far from | object |
| | overlapping with | object |
| | greater than | object |
| | smaller than | object |
| | % of image | value |

OORelation algos ready to be used

Algorithm 4 is on the left and right of

```
1: if  $X - \text{centreOfObject}_1 \leq X - \text{centreOfObject}_2$  then  
2:    $\text{Object}_1\text{IsOnTheLeft} \leftarrow \text{Object}_2$   
3: else  
4:   if  $X - \text{centreOfObject}_1 \geq X - \text{centreOfObject}_2$  then  
5:      $\text{Object}_1\text{IsOnTheRight} \leftarrow \text{Object}_2$   
6:   end if  
7: end if
```

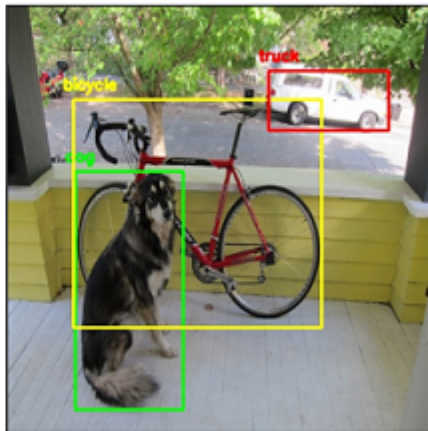
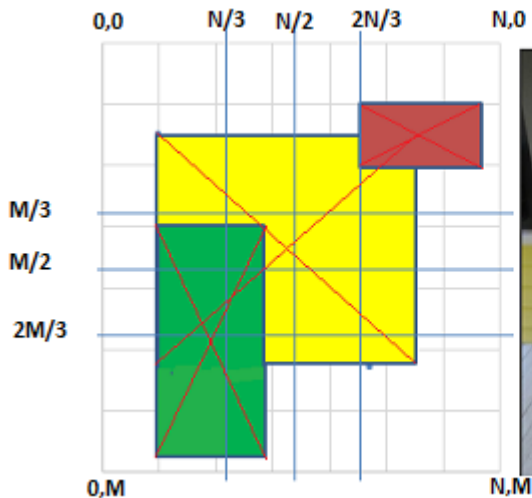
Algorithm 5 is on the top and bottom of

```
1: if  $Y - \text{centreOfObject}_1 \leq Y - \text{centreOfObject}_2$  then  
2:    $\text{Object}_1\text{IsOnTheTop} \leftarrow \text{Object}_2$   
3: else  
4:   if  $Y - \text{centreOfObject}_1 \geq Y - \text{centreOfObject}_2$  then  
5:      $\text{Object}_1\text{IsOnTheBottom} \leftarrow \text{Object}_2$   
6:   end if  
7: end if
```

Algorithm 6 close and far from

```
1: if  $\text{distance}(\text{Center1}, \text{Center2}) \leq \text{meansofdiagonalofthe2objects}$  then  
2:    $\text{Object}_1\text{IsCloseTo} \leftarrow \text{Object}_2$   
3: else  
4:   if  $\text{distance}(\text{Center1}, \text{Center2}) \geq \text{meansofdiagonalofthe2objects}$  then  
5:      $\text{Object}_1\text{IsFarFrom} \leftarrow \text{Object}_2$   
6:   end if
```

An attempt to show a graphical representation of O-O Relations in comparison to an incoming photo



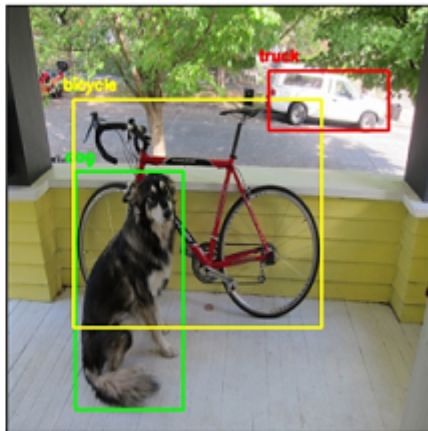
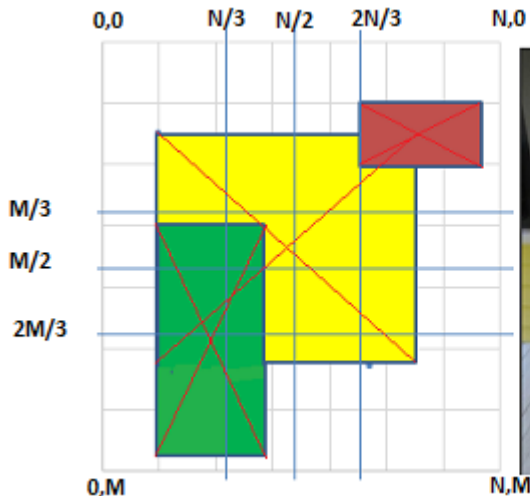
OORelation algos ready to be used, contd..

Greater and Smaller Than

Algorithm 7 greater and smaller than

```
1: if  $AreaOfObject_1 \leq AreaOfObject_2$  then  
2:    $Object_1IsSmallerThan \leftarrow Object_2$   
3: else  
4:   if  $AreaOfObject_1 \geq AreaOfObject_2$  then  
5:      $Object_1IsGreaterThen \leftarrow Object_2$   
6:   end if  
7: end if
```

An attempt to show a graphical representation of O-O Relations in comparison to an incoming photo



OORelation algos ready to be used, contd..

Overlapping With

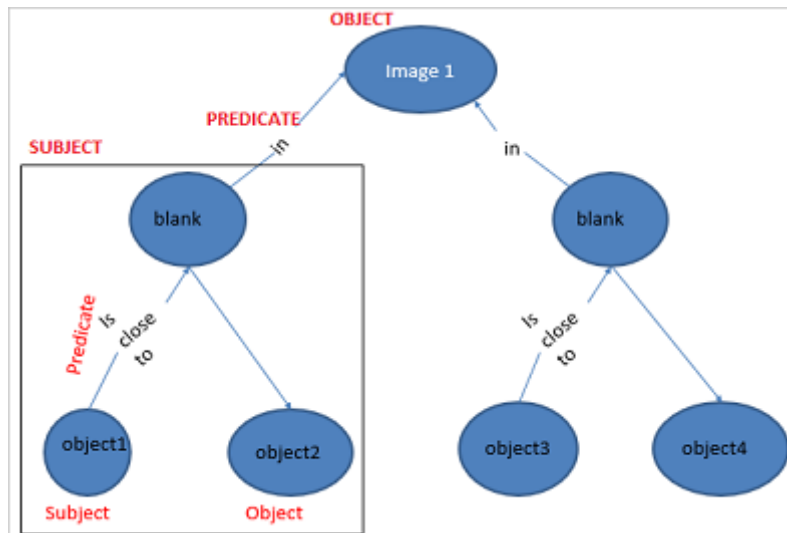
```
from shapely.geometry import Polygon
p1 = Polygon([(0,0), (1,1), (1,0)])
p2 = Polygon([(0,1), (1,0), (1,1)])
print(p1.intersects(p2))
```

True

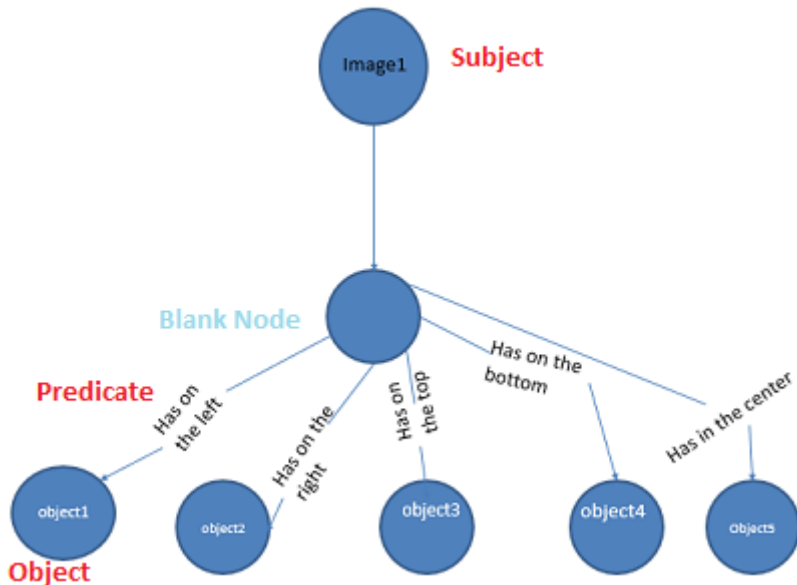
Object-Object Relation

| Object | relation | property value |
|--------|------------------|----------------|
| | left of | object |
| | right of | object |
| | top of | object |
| | bottom of | object |
| | close to | object |
| | far from | object |
| | overlapping with | object |
| | greater than | object |
| | smaller than | object |
| | % of image | value |

Reified Triple



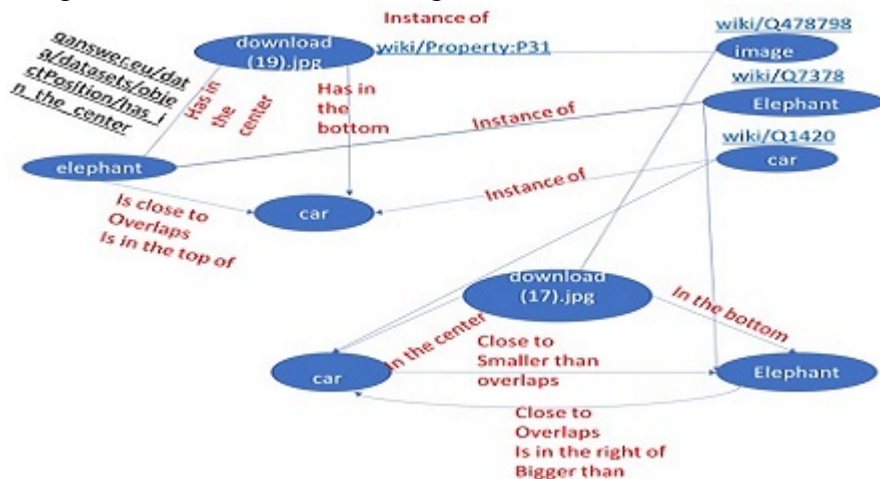
Regular Triple



Knowledge Graph

Image-Object Relation with a hint of Object-Object Relation

- Design of a semantic web modelling for extracted data.



- ① We convert the CSV file created for Image-Object Relations into
 - a RDF file
 - with the help of a Java Program
- ② and upload the RDF file into QAnswer

Results

QAnswer: images with bicycle in the right

bicycle in the right

Go

Confidence : 59 %

Is this the right answer? ☐ Yes ☐ No

/ has on the right / bicycle

LIST

IMAGES



images in the center

- QAnswer: airplane in the center

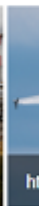
Confidence :  55 %

Is this the right answer? ☐ Yes ☐ No

/ has in the center / airplane

LIST

IMAGES



images in the left

train in the left

```
https://qanswer-frontend.univ-st-etienne.fr/user/query?  
kb=onto&user=anindamaulik
```


Qanswer: Image-Object Relation-contains

- person and chair

person and chair

Confidence : 57 %

SPARQL LIST

DID YOU MEAN

DIRECT ANSWER

Is this the right answer? ☐ Yes ☐ No

/ has on the bottom, has on the left, has on the right / human
/ has on the bottom, has on the left, has on the right / chairperson

https://upload.wikimedia.org/wikipedia/commons/e/eb/CEE_Spi



special Wikimediacommons api give images and human hand-annotated structured data

The Api

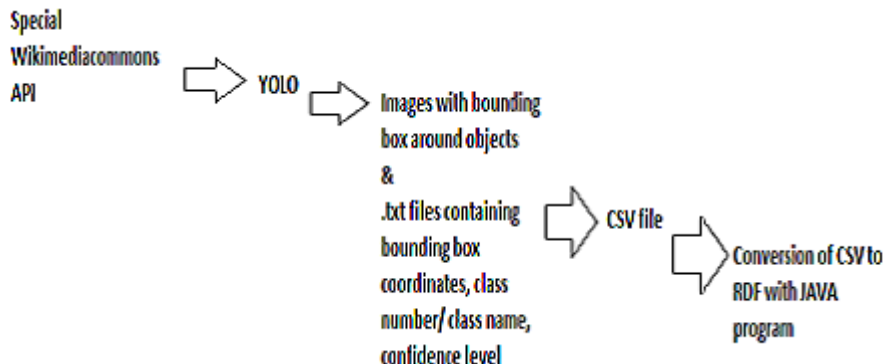
```
https://commons.wikimedia.org/w/api.php?action=query&list=search&srsearch=haswbstatement:  
P180=Q7378&srnamespace=6&format=json
```

- give images in form of json data
- and human hand-annotated structured data for some images
- Examples

- 1 what are the items portrayed in the image
- 2 who is the creator of the file
- 3 what is the source of the file
- 4 inception date of the file
- 5 copyright status
- 6 The url showing structured data

```
https://commons.wikimedia.org/wiki/File:  
African_elephants,_Lake_St_Lucia_06.jpg
```

Automation pipeline



Limitations and Future Work

Query for images on the top

Issue with confidence of detection by YOLO

/ has on the top / clock

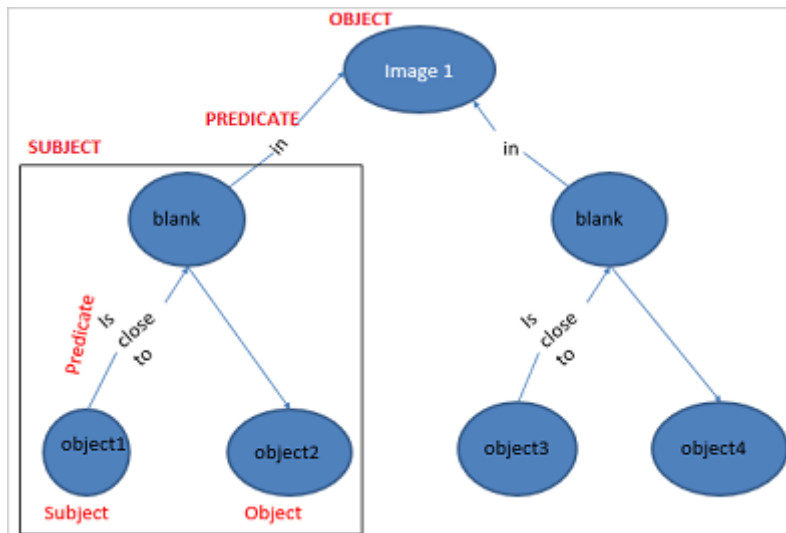
https://upload.wikimedia.org/wikipedia/commons/e/e7/Taipei_Rheinland-Office-Building-02.jpg



YOLO got overconfident



Reified Triple



Reified Triple not being interpreted by QAnswer

give me pictures with a person on right of a car

Go

Confidence :  39 %

SPARQL LIST

Is this the right answer? ☒ Yes ☐ No

/ has on the right / human

LIST

IMAGES



special Wikimediacommons api give images and human hand annotated structured data

Our time constraints

```
https://commons.wikimedia.org/w/api.php?action=query&list=
search&srsearch=haswbstatement:
P180=Q7378&srnamespace=6&format=json
```

- We are yet to handle the incoming RDF data from the api
- and merge with our RDF file,
- in order to be able to query
- "Give me pictures of bicycles from February 2019"

Conclusion

- ① We have worked on improving image search engines
- ② Computer Vision techniques is able to identify objects in images.
- ③ Semantic Web techniques give a semantic representation of the images that can be queried QAnswer.
- ④ This work can be easily used by any search or query engine to give results
- ⑤ We do not have a bench mark as of yet

- Thank you for your attention, time and patience
- Please ask me any question that you have.
- Please provide your suggestion which can be used to make my work better