

For Indian geographical database, it was created a database from the following government databases.

<http://cwc.gov.in/get-hydrological-data>

https://bhuvan.nrsc.gov.in/bhuvan_links.php

<https://www.iirs.gov.in/geoweb-services>

Sample natural language queries and their corresponding POS Tag and Chunked output for an Indian Database

1. How long is the Ganga river?

How/WRB long/RB is/VBZ the/DT Ganga/NNP river/NN ?/.

[ADVP How/WRB] long/RB [VP is/VBZ] [NP the/DT Ganga/NNP river/NN] ?/.

2. How many cities are in kerala?

How/WRB many/JJ cities/NNS are/VBP in/IN kerala/NN ?/.

[ADVP How/WRB] [NP many/JJ cities/NNS] [VP are/VBP] [PP in/IN] [NP kerala/NN] ?/.

3. Give the capital of state kerala?

Give/VB the/DT capital/NN of/IN state/NN kerala/NN ?/.

[VP Give/VB] [NP the/DT capital/NN] [PP of/IN] [NP state/NN kerala/NN] ?/.

4. Give the state name and count of rivers in each states of country India?

Give/VB the/DT state/NN name/NN and/CC count/NN of/IN rivers/NNS in/IN each/DT states/NNS of/IN country/NN India/NNP ?/.

[VP give/VB] [NP the/DT state/NN name/NN and/CC count/NN] [PP of/IN] [NP rivers/NNS]

[PP in/IN] [NP each/DT states/NNS] [PP of/IN] [NP country/NN India/NNP] ?/.

5. Give the total population of Mumbai?

Give/VB the/DT total/JJ population/NN of/IN Mumbai/NNP ?/.

[VP Give/VB] [NP the/DT total/JJ population/NN] [PP of/IN] [NP Mumbai/NNP] ?/.

6. Give the state name having highest population in country India?

Give/VB the/DT state/NN name/NN having/VBG highest/JJS population/NN in/IN country/NN

India/NNP ?/.

[VP Give/VB] [NP the/DT state/NN name/NN] [VP having/VBG] [NP highest/JJS population/NN] [PP in/IN] [NP country/NN India/NNP] ?/.

7. Give the name of rivers in state Kerala?

Give/VB the/DT name/NN of/IN rivers/NNS in/IN state/NN Kerala/NNP ?/.

[VP Give/VB] [NP the/DT name/NN] [PP of/IN] [NP rivers/NNS] [PP in/IN] [NP state/NN

Kerala/NNP] ?/.

8. Give the state name where city trivandrum is located?

Give/VB the/DT state/NN name/NN where/WRB city/NN trivandrum/NN is/VBZ located/VBN

?/.

Attachment to manuscript [Click here to download Attachment to manuscript NLP QUERY.doc](#)

chunked output

[VP Give/VB] [NP the/DT state/NN name/NN] [ADVP where/WRB] [NP city/NN trivandrum/NN] [VP is/VBZ located/VBN] ?/.

9. Which is the largest river in India?

Which/WDT is/VBZ the/DT largest/JJS river/NN in/IN India/NNP ?/.

chunked output

[NP Which/WDT] [VP is/VBZ] [NP the/DT largest/JJS river/NN] [PP in/IN] [NP India/NNP] ?/.

10. Give the count of rivers in each states of country India?

Give/VB the/DT count/NN of/IN rivers/NNS in/IN each/DT states/NNS of/IN country/NN India/NNP ?/.

[VP Give/VB] [NP the/DT count/NN] [PP of/IN] [NP rivers/NNS] [PP in/IN] [NP each/DT states/NNS] [PP of/IN] [NP country/NN India/NNP] ?/.

11. Tell me names and addresses of our suppliers with the telephone number start with 402?

Tell/VB me/PRP names/NNS and/CC addresses/NNS of/IN our/PRP\$ suppliers/NNS with/IN

the/DT telephone/NN number/NN start/VB with/IN 402/CD ?/.

[VP Tell/VB] [NP me/PRP] [NP names/NNS and/CC addresses/NNS] [PP of/IN] [NP our/PRP\$ suppliers/NNS] [PP with/IN] [NP the/DT telephone/NN number/NN] [VP start/VB] [PP with/IN] [NP 402/CD] ?/.

12. Tell me the number of registered candidates in the course Computer Science and Engineering during the year 2014?

Tell/VB me/PRP the/DT number/NN of/IN registered/JJ candidates/NNS in/IN the/DT course/NN Computer/NNP Science/NNP and/CC Engineering/NNP during/IN the/DT year/NN 2014/CD ?/.

[VP Tell/VB] [NP me/PRP] [NP the/DT number/NN] [PP of/IN] [NP registered/JJ candidates/NNS] [PP in/IN] [NP the/DT course/NN Computer/NNP Science/NNP and/CC Engineering/NNP] [PP during/IN] [NP the/DT year/NN 2014/CD] ?/.

13. How many students have registered for Machine learning elective from Computer Science and Electronics branch in the year 2016?

How/WRB many/JJ students/NNS have/VBP registered/VBN for/IN Machine/NN learning/VBG elective/NN from/IN Computer/NNP Science/NNP and/CC Electronics/NNP branch/NN in/IN the/DT year/NN 2016/CD ?/.

[ADVP How/WRB] [NP many/JJ students/NNS] [VP have/VBP registered/VBN] [PP for/IN] [NP Machine/NN learning/VBG elective/NN] [PP from/IN] [NP Computer/NNP Science/NNP and/CC Electronics/NNP branch/NN] [PP in/IN] [NP the/DT year/NN 2016/CD] ?/.

14. What is the class average for Maths in first semester Computer Science?

What/WP is/VBZ the/DT class/NN average/NN for/IN Maths/NNS in/IN first/JJ

semester/NN Computer/NNP Science/NNP ?/.

[NP What/WP] [VP is/VBZ] [NP the/DT class/NN average/NN] [PP for/IN] [NP Maths/NNS] [PP in/IN] [NP first/JJ semester/NN Computer/NNP Science/NNP] ?/.

15. What is the name and product id of the products that have been returned to the supplier due to manufacturing defect?

What/WP is/VBZ the/DT name/NN and/CC product/NN id/NN of/IN the/DT products/NNS that/WDT have/VBP been/VBN returned/VBN to/TO the/DT supplier/NN due/JJ to/TO manufacturing/VBG defect/NN ?/.

[NP What/WP] [VP is/VBZ] [NP the/DT name/NN and/CC product/NN id/NN] [PP of/IN] [NP the/DT products/NNS] [NP that/WDT] [VP have/VBP been/VBN returned/VBN] [PP to/TO] [NP the/DT supplier/NN] [ADJP due/JJ] [PP to/TO] [NP manufacturing/VBG defect/NN] ?/.

16. Give the list of suppliers who have delivered maximum number of products?

Give/VB the/DT list/NN of/IN suppliers/NNS who/WP have/VBP delivered/VBN maximum/JJ number/NN of/IN products/NNS ?/.

[VP Give/VB] [NP the/DT list/NN] [PP of/IN] [NP suppliers/NNS] [NP who/WP] [VP have/VBP delivered/VBN] [NP maximum/JJ number/NN] [PP of/IN] [NP products/NNS] ?/.

Consider the query: “Give the state name and count of rivers in each states of country India?”
POS tagged output: Give/VB the/DT state/NN name/NN and/CC count/NN of/IN rivers/NNS in/IN each/DT states/NNS of/IN country/NN India/NNP ?/

Chunked output: [VP give/VB] [NP the/DT state/NN name/NN and/CC count/NN] [PP of/IN] [NP rivers/NNS] [PP in/IN] [NP each/DT states/NNS] [PP of/IN] [NP country/NN India/NNP] ?/

As per the query translation algorithm, the state name and count is having tag NN, so these need

to be displayed in output. But count is an aggregate function, as per the query template we need

to use group by statement. To find the column names for the group by statement. As *state_name*

to display, one of the group by column name is *state_name* and other is *river_name* for displaying the count of rivers. Then rivers and states are having NN tag, the proposed method sets these tags as table names. The country tag NN is followed by NNP tag India, it is set as constraint for WHERE class. These details are passed through query generator template. The standard query generator template would be:

```
SELECT $column_name_list
FROM $table_name_list
WHERE (if any) $constraint_list
GROUP BY (if any) $group_by_list
HAVING (if any) $having_list
ORDER BY $order_list
```

Where *column_name_list*, *table_name_list*, *constraint_list*, *group_by_list*, *having_list*, *order_list*

are variable. Value of these variables is generated by the query translation algorithm. In our example, *column_name_list* is having values *state_name* and *count(*)*. Rivers and states are the values stored on *table_name_list*. Proper alias is generated if multiple table names in this list. *Constraint_list* is having the values *r.state_name=s.state_name* and *s.country_name='India'* and *group_by_list* have the values *river_name* and *state_name*. So the resultant query would like

```
SELECT state_name,count(*) FROM rivers r,states s
      WHERE r.state_name=s.state_name
            AND s.country_name='India'
      GROUP BY river_name,state_name;
```