THE-QA (Technical, Hard and Explainable Question Answering)

Akshar Patel
adpate11@asu.edu

Karthiga Hariharan
kharihar@asu.edu

Shahabudeen Sajith
ssajith@asu.edu

Shuddhatm Jain
shuddhatm@asu.edu

Spoorthi Karnati
skarnat1@asu.edu

1 Problem Definition

The Artificial Intelligence Systems need to understand visual and textual inputs. Combination of those two inputs are required to instruct and explain. Lot of research is going on visual understanding and textual understanding in isolation. An intelligent robot might learn about its tasks and environment by observing both language and gesture. Hence more research has to be done on combined version of visual and textual understanding. THE-QA is the project which develops framework for answering hard questions about combined visual and textual inputs, and provides supporting explanations. The challenge here is that we need to build a system that needs to generate answer by employing an integrated approach of deep model-based visual recognition and natural language processing, and knowledge representation. This project involves data set creation and baseline model formation.

2 Examples

2.1 Example 1

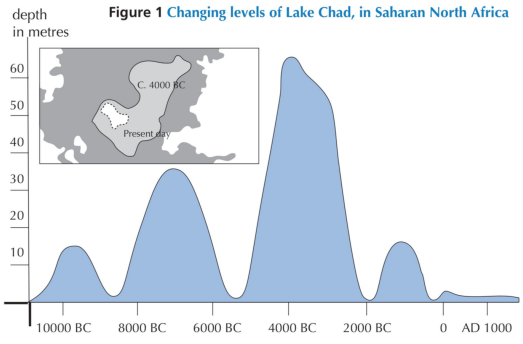


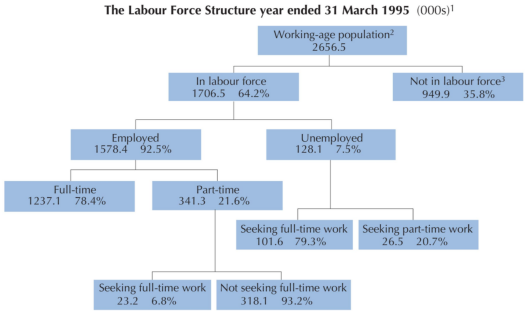
Figure 1 Changing levels of Lake Chad, in Saharan North Africa

Figure 1 shows changing levels of Lake Chad, in Saharan North Africa. Lake Chad disappeared completely in about 20,000 BC, during the last Ice Age. In about 11,000 BC it reappeared. **Today**, its level is **about the same** as it was in AD 1000.

What is the depth of Lake Chad **today**?

1. **About two metres**.
2. About fifteen metres.
3. About fifty metres.
4. It has disappeared completely.
5. The information is not provided.

2.2 Example 2



The **tree** diagram in Figure 2 shows the structure of a countrys labour force or working-age population. The **total population** of the country in 1995 was about 3.4 million.

In which part, a business woman, aged 43, who works a **sixty-hour week**, would be included?

1. **Labor force "employed"**.
2. Labor force "unemployed".
3. Not in labor force.
4. Not included in any category.

2.3 Example 3

The table in Figure 3 shows the recommended Zedland **shoe sizes** corresponding to various **foot lengths**.

Marinas **feet** are 163 mm long. Use the table to determine which Zedland **shoe size** Marina should try on.

Conversion table for kids shoe sizes in Zedland

From (in mm)	To (in mm)	Shoe size
107	115	18
116	122	19
123	128	20
129	134	21
135	139	22
140	146	23
147	152	24
153	159	25
160	166	26
167	172	27
173	179	28
180	186	29
187	192	30

1. 22
2. 24
3. **26**
4. 28

3 Data Source

CIA's World Factbook (`https://www.cia.gov/library/publications/resources/the-world-factbook/`) provides information on the history, people, government, economy, geography, communications, transportation, military, and transnational issues for 267 world entities. The data can be classified into three flavors.

3.1 Time Series

A series of values of a quantity obtained at successive times, often with equal intervals between them. The following labels belong to this flavor.

1. central-bank-discount-rate
2. commercial-bank-prime-lending-rate
3. current-account-balance
4. debt-external
5. distribution-of-family-income-gini-index
6. exchange-rates
7. exports
8. gdp-per-capita-ppp
9. gdp-purchasing-power-parity
10. gdp-real-growth-rate
11. gross-national-saving
12. imports
13. inflation-rate-consumer-prices
14. market-value-of-publicly-traded-shares
15. military-expenditures
16. public-debt
17. reserves-of-foreign-exchange-and-gold
18. stock-of-broad-money
19. stock-of-direct-foreign-investment-abroad
20. stock-of-direct-foreign-investment-at-home
21. stock-of-domestic-credit
22. stock-of-narrow-money
23. unemployment-rate

3.2 Categorical

Data consisting for categorical variables (that can take on one of a limited, and usually fixed number of possible values). The following labels belong to this flavor.

1. agriculture-products
2. exports-commodities
3. imports-commodities
4. industries
5. map-references
6. natural-resources

3.3 Comparative

The item-by-item comparison of two or more comparable sets of data. The following labels belong to this flavor.

1. Age structure 55-64 years
2. Airports - with paved runways total
3. Airports - with unpaved runways total
4. Airports total
5. Area land
6. Birth rate
7. Broadband - fixed subscriptions subscriptions per 100 inhabitants
8. Budget expenditures
9. Budget surplus (+) or deficit (-)
10. Carbon dioxide emissions from consumption of energy
11. Children under the age of 5 years underweight
12. Coastline
13. Contraceptive prevalence rate
14. Crude oil - exports
15. Crude oil - imports
16. Crude oil - production
17. Crude oil - proved reserves
18. Death rate
19. Dependency ratios elderly dependency ratio
20. Education expenditures
21. Electricity - consumption
22. Electricity - exports
23. Electricity - from fossil fuels
24. Electricity - from hydroelectric plants
25. Electricity - from nuclear fuels
26. Electricity - from other renewable sources
27. Electricity - imports
28. Electricity - installed generating capacity
29. Electricity - production
30. Electricity access electrification - total population
31. GDP - composition, by end use government consumption
32. GDP - composition, by sector of origin agriculture
33. HIV/AIDS - adult prevalence rate
34. HIV/AIDS - deaths
35. HIV/AIDS - people living with HIV/AIDS
36. Health expenditures
37. Heliports
38. Hospital bed density
39. Household income or consumption by percentage share lowest 10
40. Industrial production growth rate
41. Infant mortality rate male
42. Internet users total
43. Irrigated land
44. Labor force
45. Labor force - by occupation agriculture
46. Land boundaries total
47. Land use agricultural land
48. Life expectancy at birth female
49. Literacy female
50. Maternal mortality rate
51. Median age male
52. Merchant marine total
53. National air transport system inventory of registered aircraft operated by air carriers
54. Natural gas - consumption
55. Natural gas - exports
56. Natural gas - imports

57. Natural gas - production
58. Natural gas - proved reserves
59. Net migration rate
60. Obesity - adult prevalence rate
61. Physicians density
62. Population
63. Population below poverty line
64. Population growth rate
65. Railways total
66. Refined petroleum products - consumption
67. Refined petroleum products - exports
68. Refined petroleum products - imports
69. Refined petroleum products - production
70. Roadways total
71. School life expectancy (primary to tertiary education) male
72. Sex ratio 15-24 years
73. Taxes and other revenues
74. Telephones - fixed lines total subscriptions
75. Telephones - mobile cellular subscriptions per 100 inhabitants
76. Total fertility rate
77. Unemployment, youth ages 15-24 female
78. Urbanization urban population
79. Waterways

4 Scraping

Using Scrapy (`https://scrapy.org/`), a Python framework to extract data from websites, the above data was compiled in `json` format.

4.1 Time Series

```
{
  "United States": {
    "_x_": [],
    "military-expenditures": [
      ["2016", "3.29% of GDP"],
      ["2015", "3.3% of GDP"],
      ["2014", "3.51% of GDP"],
      ["2013", "3.83% of GDP"],
      ["2012", "4.24% of GDP"]
    ],
    "_y_": []
  }
}
```

4.2 Categorical

```
{
  "Canada": {
    "_x_": [],
    "agriculture-products": [
      "wheat", "barley", "tobacco",
      "fruits", "vegetables", "fish"
    ],
    "_y_": []
  }
}
```

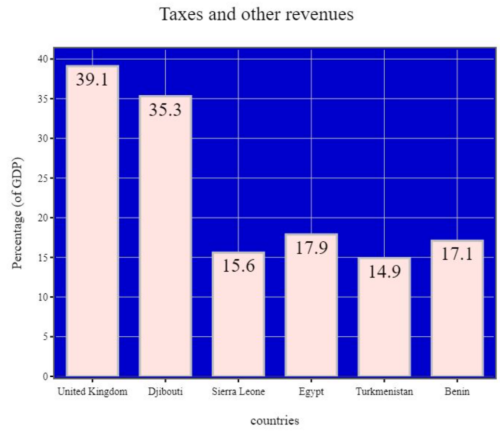
4.3 Comparative

```
{
  "Population": {
    "_x_": "",
    "United States": "329,256,465",
    "Canada": "35,881,659",
    "Mexico": "125,959,205",
    "_y_": ""
  }
}
```

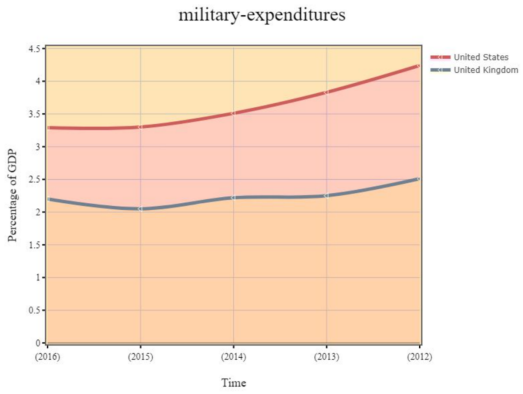
5 Figures

Using Plotly (`https://github.com/plotly/plotly.py`), an open-source, graphing library for Python the above `json` data can be converted to different graphical figures like the ones shown below.

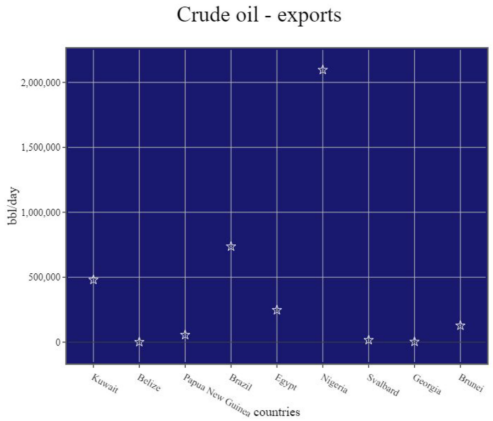
5.1 Bar



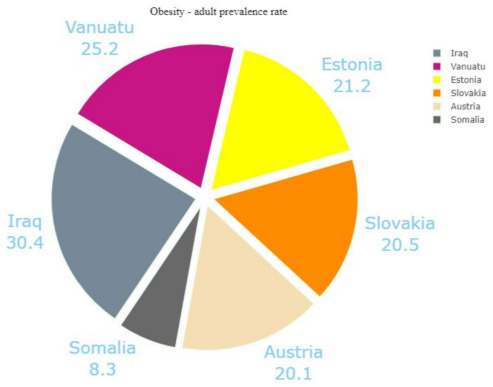
5.2 Line/Area



5.3 Scatter



5.4 Pie



6 Common Sense Knowledge

Generating the paragraph and the accompanying question would involve injecting some kind of common sense knowledge. Since the figures depict graphical data, the question phrase will generally contain quantitative/temporal terms. Synonyms to these terms can introduce variety in the questions that are generated. A template-based question generation will work fairly well.

Common sense knowledge on the other hand will be tightly coupled to the data attribute that is depicted. For example birth/death rate can be related to population; GDP can be related to purchasing power etc. These relations need to be embedded manually in the paragraph. The accompanying questions can then be linked to this knowledge.

7 Sample Questions

7.1 Figure 1

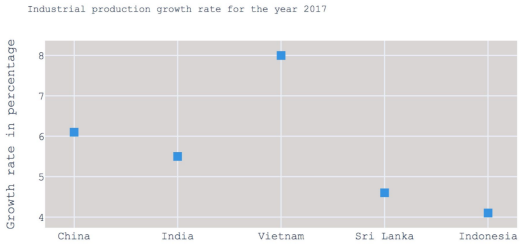


Figure 1 shows the industrial production growth rate for various countries in the year of 2017. Industrial production figures are used by central banks to measure inflation, as high levels of industrial production can lead to uncontrolled levels of consumption and rapid inflation. In 2017,

the industrial production growth rate of India was similar to that of Pakistan.

7.1.1 Question 1

Which of these countries is most likely to have a rapid inflation when compared to Pakistan?

1. India
2. **Vietnam**
3. Sri Lanka
4. Indonesia

7.1.2 Question 2

The industrial production growth rate of Pakistan is almost equal to

1. 7.5
2. 6.5
3. **5.5**
4. 4.5

7.1.3 Question 3

Consider all the given 6 countries. The industrial production growth rate of Sri Lanka is less than how many countries

1. 1
2. 2
3. 3
4. **4**

7.2 Figure 2

	Canada	Argentina	Germany	India
Wheat	Y	Y	Y	Y
Barley	Y	N	Y	N
Tea	N	Y	N	Y
Potatoes	N	N	Y	Y
Fish	Y	Y	N	Y
Tobacco	Y	Y	N	N

The agricultural produce of various countries is depicted in Figure 2. Y indicates Yes, N indicates No.

7.2.1 Question 1

Which country is most likely to import Fish?

1. Canada
2. Argentina
3. **Germany**
4. India

7.2.2 Question 2

Tea, Potatoes and Coffee thrive well in tropical climate. Which country produces coffee?

1. Canada
2. Argentina
3. Germany
4. **India**

7.2.3 Question 3

Countries which produce fish have a coastline. Which country is landlocked?

1. Canada
2. Argentina
3. **Germany**
4. India

8 Future Work

1. Identify common sense relations for the data attributes scraped from CIA website.
2. Generate a data-set (figure, paragraph, question, options, correct option, reasoning) of 1000 samples.
3. Baseline model evaluation.