A decorative graphic on the left side of the slide consisting of two overlapping parallelograms. The front one is blue and the back one is a light greenish-blue. They are both tilted at an angle.

# Building end-to-end QA model with reward based objective

TEAM REINFORCE (Prashant and Eti)



# Exploring Data: SearchQA

- 140k+ QA pairs, 6.9M snippets
- 49.6 snippets on average
- Each snippet is 37.3 tokens long on average
- Answers are on average 1.47 tokens long
- 1.2M unique tokens
- Additional Meta-Data:
  - Category, Dollar value, Show number and Air date
  - From Google: URL, Title and a set of related links (often none) for each snippet

```
{
  "category": "HISTORY",
  "air_date": "2004-12-31",
  "question": "'For the last 8 years of his life, Galileo was under house arrest for espousing this man's theory'",
  "value": "$200",
  "answer": "Copernicus",
  "round": "Jeopardy!",
  "show_number": "4680"
},
```



# Exploring Data: WikiSuggest

- # of unique queries, # of examples: 3.47M
- # of words / query: 5.03
- # of tokens / doc: 5962.2
- % answer string exists: 100
- avg # of answer match: 13.95
- % match first sent: 33.6

$s_1$ : The 2011 Joplin tornado was a catastrophic EF5-rated multiple-vortex tornado that struck Joplin, Missouri . . .

$d$ :  $s_4$ : It was the third tornado to strike Joplin since May 1971.

$s_5$ : Overall, the tornado killed **158 people** . . . , injured some 1,150 others, and caused damages . . .

$x$ : how many people died in joplin mo tornado

$y$ : 158 people



# Vocabulary

Some options:

- Use pre-trained embeddings (eg. Glove)
- Use pre-trained embeddings (which were trained on a similar task/dataset)
- Use word-2-vec to do unsupervised training of word embeddings on these datasets



# Literature Review

Bi-DAF for Machine Comprehension (Seo et al 2017):

- Word + Char embeddings passed to an LSTM to get representations of context and query words
- Query2Context: Signifies which context words are most similar to one of the query words and hence important for answering the query
- Context2Query: Calculates relevance of each query word with respect to each context word
- Finally, the hidden representations and the two module scores are fused and inputted to an LSTM network (Modelling layer)
- Span selection in final output layer

# Literature Review

## Bi-directional Attention Flow Demo for [Stanford Question Answering Dataset \(SQuAD\)](#)

**Direction** : Select a paragraph and write your own question. The answer is always a subphrase of the paragraph - remember it when you ask a question!

Select Paragraph

[03] Nikola\_Tesla

Paragraph

Nikola Tesla (Serbian Cyrillic: Никола Тесла; 10 July 1856 – 7 January 1943) was a Serbian American inventor, electrical engineer, mechanical engineer, physicist, and futurist best known for his contributions to the design of the modern alternating current (AC) electricity supply system.

Question

What does AC stand for?

new question!

Answer

modern alternating current

Reference : [Minjoon Seo](#), [Aniruddha Kembhavi](#), [Ali Farhadi](#), [Hannaneh Hajishirzi](#). "Bidirectional Attention Flow for Machine Comprehension" [\[link\]](#)

Demo by : [Sewon Min](#)

Ref: <https://allenai.github.io/bi-att-flow/>

# Bi-DAF Failure Cases

## Select Paragraph

[03] Nikola\_Tesla ▼

## Paragraph

Nikola Tesla (Serbian Cyrillic: Никола Тесла; 10 July 1856 – 7 January 1943) was a Serbian American inventor, electrical engineer, mechanical engineer, physicist, and futurist best known for his contributions to the design of the modern alternating current (AC) electricity supply system.

## Question

Write own question ▼

Who was Nikola ?

new question!

## Answer

Nikola Tesla

## Select Paragraph

[03] Nikola\_Tesla ▼

## Paragraph

Nikola Tesla (Serbian Cyrillic: Никола Тесла; 10 July 1856 – 7 January 1943) was a Serbian American inventor, electrical engineer, mechanical engineer, physicist, and futurist best known for his contributions to the design of the modern alternating current (AC) electricity supply system.

## Question

Write own question ▼

who was Nikola Tesla ?

new question!

## Answer

Serbian American inventor

# Bi-DAF Failure Cases

## Select Paragraph

[01] Warsaw

## Paragraph

One of the most famous people born in Warsaw was Maria Skłodowska-Curie, who achieved international recognition for her research on radioactivity and was the first female recipient of the Nobel Prize. Famous musicians include Władysław Szpilman and Frédéric Chopin. Though Chopin was born in the village of Żelazowa Wola, about 60 km (37 mi) from Warsaw, he moved to the city with his family when he was seven months old. Casimir Pulaski, a Polish general and hero of the American Revolutionary War, was born here in 1745.

## Question

Write own question

who was Władysław Szpilman ?

new question!

## Answer

Frédéric Chopin

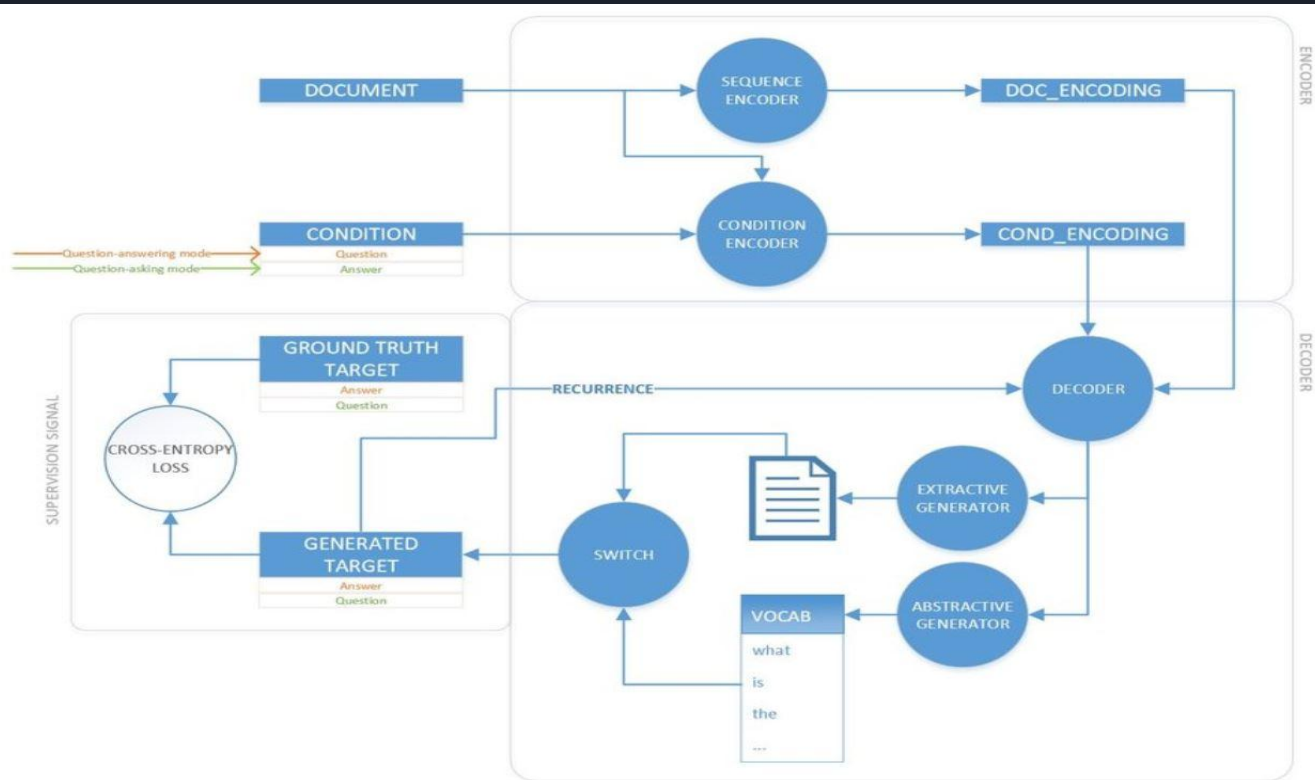




# Literature Review

## A Joint Model for Question Answering and Question Generation ( Wang et al. 2017)

- Encoder :
  - A word is represented as a concatenation of word and character embeddings
  - They are encoded with another BiLSTM into annotation vectors  $h^d_i$  and  $h^c_j$  (for the document and the condition sequence, respectively)
  - To better encode the condition, the encodings of the document words that appear in the condition sequence are also fed to the conditional encoder.
- Decoder :
  - At each generation step, the decoder decides adaptively whether (a) to generate from a decoder vocabulary or (b) to point to a word in the source sequence (and copy over).
  - Uses attention over entire document to improve the performance of the decoder





# Next Steps (for the upcoming week)

Prashant:

- Implement the baseline Bi-DAF model
- Use pre-trained Glove embeddings to train the model
- Get baseline metric scores on both the datasets
- Do preliminary error analysis to figure out the shortcomings of the baseline model and common error types

Eti:

- Implement the Joint Model
- Get baseline metric scores on both the datasets
- Do preliminary error analysis to figure out the shortcomings of the baseline model and common error types