

# Question Generation

A Literature Review

Baozheng Li

Qichao Jiang

Zimo Wang

# Our Work

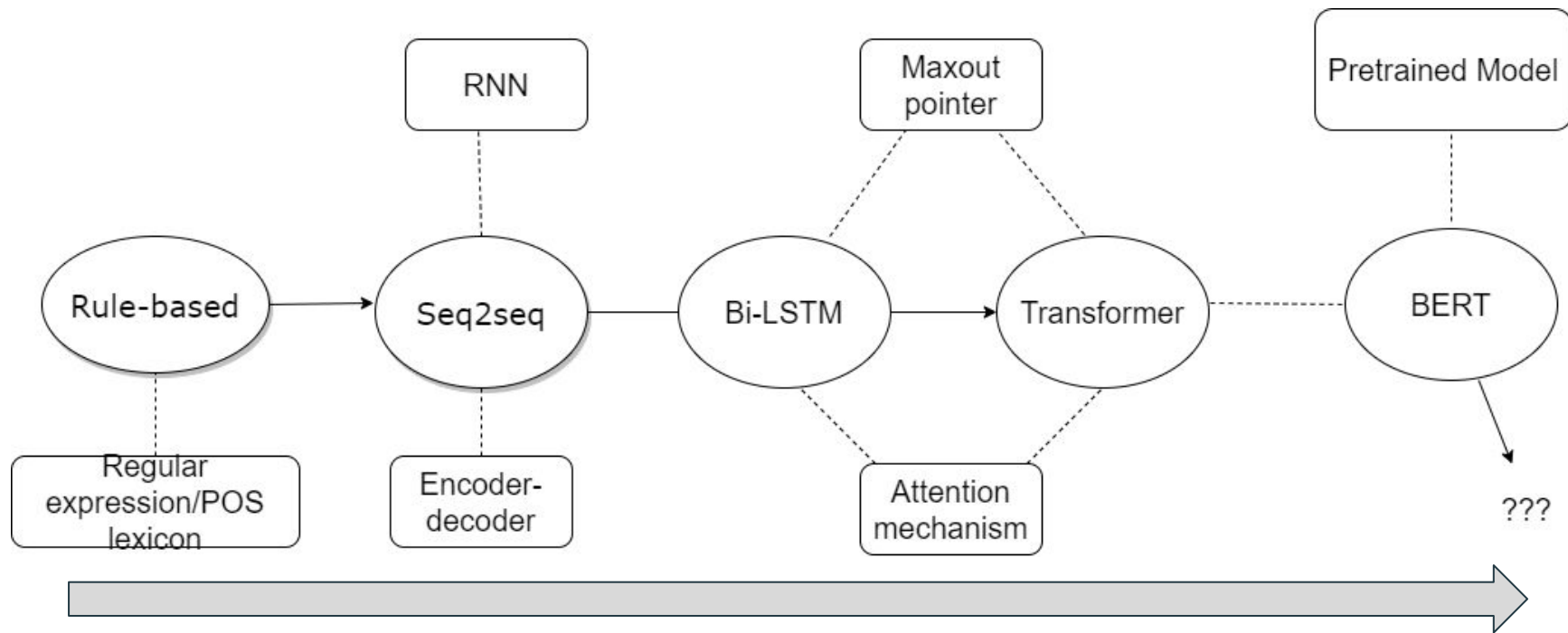
→ 32 Papers published between 2018 to 2022

→ Current Work Progress

→ Review Structure

- ◆ Abstract
- ◆ Introduction
- ◆ Summary of Previous Reviews
- ◆ Review Objective
- ◆ Review Contents
  - Architectures
  - Preprocessing Methods
  - Encoder-Decoder Models
  - Datasets
  - Metrics
- ◆ Limitations of current work
- ◆ Conclsusions and Future Work

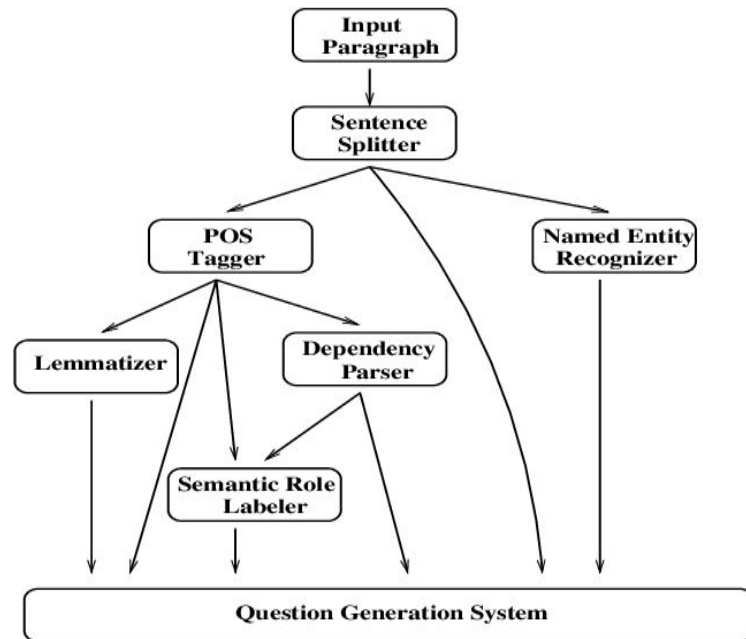
# Trends in Question Generation - Models



Previous Question Generation techniques

# Preprocessing methods

- Standard NLP
- QG-Specific
  - Sentence simplification
  - Sentence classification
  - Content selection



# Datasets

## Quiz-Style Question Generation for News Stories

Adam D. Lelkes\*  
Google Research  
lelkes@google.com

Vinh Q. Tran\*  
Google Research  
vqtran@google.com

Cong Yu  
Google Research  
congyu@google.com

Dataset	mid-training	Domain	Type	QA Pairs	Avg. pairs per passage	Avg. question length	Avg. answer length
SQuAD		Wikipedia	Reading comprehension	98,169	4.68	59.6	20.1
Natural Questions		Wikipedia	Search query	99,486	1.55	48.3	24.9
NewsQA		News	Reading comprehension	90,529	7.50	37.0	25.5
NewsQuizQA		News	Pop quiz	20,024	4.00	77.8	27.1

Prefix the input to differentiate the style of datasets:

"Style SQuAD:"

"Style NQ:"

"Style NewsQA:"

# Metrics

Gold	Generated	WER	BLEU-Score(1gram)
where was PERSON 2 born?	what is the birth place of PERSON 2?	4	$3/9 = 0.33$
where was PERSON 2 born?	when was PERSON 2 born?	1	$5/6 = 0.83$

## References & Picture credit

[https://www.researchgate.net/figure/Preprocessing-Flow-Chart\\_fig1\\_228944445](https://www.researchgate.net/figure/Preprocessing-Flow-Chart_fig1_228944445)

<https://medium.com/@diksha19078/automatic-question-generation-system-fe012fa040cb>

<https://link.springer.com/article/10.1007/s40593-019-00186-y#Sec2>

Question time...

