

A Supplementary Materials

A.1 Example Question-Answer Pairs from the Corpus

We provide more examples of QA pairs from the corpus, the red answer spans correspond to the questions in order.

Paragraph 2 France has long been a global centre of art, science, and philosophy. It hosts Europe's fourth-largest number of cultural UNESCO World Heritage Sites and receives around **83 million** foreign tourists annually, the most of any country in the world. France is a developed country with the world's sixth-largest economy by **nominal GDP** and ninth-largest by purchasing power parity. In terms of aggregate household wealth, it ranks fourth in the world. France performs well in international rankings of education, health care, life expectancy, and human development. France remains a great power in the world, being a founding member of the United Nations, where it serves as one of the five permanent members of the UN Security Council, and a founding and leading member state of the European Union (EU). It is also a member of the Group of 7, **North Atlantic Treaty Organization** (NATO), Organisation for Economic Co-operation and Development (OECD), the World Trade Organization (WTO), and La Francophonie.

Questions:

Q1: how many foreign tourists does france have ?

Q2: what is france 's sixth-largest economy ?

Q3: what does nato stand for ?

Paragraph 2 The United States embarked on a vigorous expansion across North America throughout the 19th century, displacing American Indian tribes, acquiring new territories, and gradually admitting new states until it spanned the continent by **1848**. During the second half of the 19th century, the **American Civil War** led to the end of legal slavery in the country. By the end of that century, the United States extended into the Pacific Ocean, and its economy, driven in large part by the Industrial Revolution, began to soar. The **Spanish-American War** and confirmed the country's status as a global military power. The United States emerged from as a global superpower, the first country to develop nuclear weapons, the only country to use them in warfare, and a permanent member of the United Nations Security Council. It is a founding member of the Organization of American States (OAS) and various other Pan-American and international organisations. The end of the Cold War and the dissolution of the Soviet Union in 1991 left the United States as the world's sole superpower.

Questions:

Q1: in what year did the united states span the continent ?

Q2: what war led to the end of legal slavery ?

Q3: what was the name of the war that confirmed the country ?

Paragraph 3 The **International Standard Name Identifier** (ISNI) is an identifier for uniquely identifying the public identities of contributors to media content such as books, TV programmes, and newspaper articles. Such an identifier consists of **16 digits**. It can optionally be displayed as divided into **four blocks**.

Questions:

Q1: what is an example of a identifier name ?

Q2: how many digits does an identifier have ?

Q3: how long can the identifier be displayed ?

Paragraph 4 India, officially the **Republic of India**, is a country in **South Asia**. It is the seventh-largest country by area, the second-most populous country (with over **1.2 billion** people), and the most populous democracy in the world. It is bounded by the Indian Ocean on the south, the Arabian Sea on the southwest, and the Bay of Bengal on the southeast. It shares land borders with Pakistan to the west; China, Nepal, and Bhutan to the northeast; and Myanmar (Burma) and Bangladesh to the east. In the Indian Ocean, India is in the vicinity of **Sri Lanka and the Maldives**. India's Andaman and Nicobar Islands share a maritime border with **Thailand and Indonesia**. Its capital is **New Delhi**; other metropolises include Mumbai, Kolkata, Chennai, Bangalore, Hyderabad and Ahmedabad.

Questions:

Q1: what is india 's country called ?

Q2: where is the republic of india located ?

Q3: how many people are in the second-most populous country ?

Q4: what is the vicinity of india ?

Q5: which two countries did the nicobar islands play a maritime border with ?

Q6: what is the capital of india ?

A.2 Human Rater Guidelines

We provide the following guidelines for the raters,

Categories	Rating scheme
grammaticality	Given only the question itself, is it grammatical?
making sense	Given just the question and the surrounding context in the passage, does the question make sense?
answerability	Given just the question and the surrounding context in the passage and the answer (and regardless of “Grammaticality” and “Making Sense”), is the question answerable by the corresponding answer span?

Table 1: Guidelines for the raters. For each category, the human raters are required to give a rating ranging from 1 to 5 (5 = fully satisfying the rating scheme, 1 = completely not satisfying the rating scheme, 3 = the borderline cases.)

A.3 Training and Implementation Details

For the question generation model, the input and output vocabularies are collected from the training data, we keep the 50k most frequent words. The size of word embedding and LSTM hidden states are set to 128 and 256, respectively. We use dropout (Srivastava et al., 2014) with probability $p = 0.3$. The model parameters are initialized randomly using a uniform distribution between $(-0.1, 0.1)$. We use Stochastic Gradient Descent (SGD) as optimization algorithm with a mini-batch size 64. We also apply gradient clipping (Pascanu et al., 2013) with range $(-5, 5)$ during training. The best models are selected based on the perplexity (lowest) on the development set. In all experiments, we use the same split of Du et al. (2017) of SQuAD dataset into training, development and test sets. We use beam search during decoding to get better results. We set the beam size to 3 in the experiments and corpus generation.

References

- Xinya Du, Junru Shao, and Claire Cardie. 2017. Learning to ask: Neural question generation for reading comprehension. In *Association for Computational Linguistics (ACL)*.
- Razvan Pascanu, Tomas Mikolov, and Yoshua Bengio. 2013. On the difficulty of training recurrent neural networks. In *International Conference on Machine Learning*. pages 1310–1318.
- Nitish Srivastava, Geoffrey Hinton, Alex Krizhevsky, Ilya Sutskever, and Ruslan Salakhutdinov. 2014. Dropout: A simple way to prevent neural networks from overfitting. *The Journal of Machine Learning Research* 15(1):1929–1958.