

Annotated Question and Answer Dataset for Security Export Control

Akihiko Obayashi¹, Rafal Rzepka^{2*}

¹Center for Innovation and Business Promotion,
Hokkaido University, Kita-ku, North 21, West 10, Sapporo, Japan

²Faculty of Information Science and Technology,
Hokkaido University, Kita-ku, North 14, West 8, Sapporo, Japan
obayashi@mcip.hokudai.ac.jp, rzepka@ist.hokudai.ac.jp

Abstract

This paper introduces a set of questions and answers in Japanese language for the topics related to security export control. Unlike the most available datasets for question answering our set comprises of very detailed expert knowledge in both queries and replies. The knowledge is not widely shared which makes it difficult to simply apply current neural approaches and its small size limits fine tuning. By introducing this data we count on increasing number of researchers extending contemporary NLP methods to be applicable to very precise expert systems. As the queries may often require additional questions for clarification, this dataset can also be utilized for testing sophisticated task-oriented dialog systems.

1 Introduction

1.1 About Security Export Control

Security Export Control is to control transfer of technologies and export of goods for the purpose of preserving the peace and security of the international community. It works to prevent transfer of the technologies and goods that can be potentially diverted to weapons or military use to any such person who might conduct activities of concern as a nation or terrorists who could threaten the peace and security of the international community. The only existing support system for determining whether transfer of technologies or export of goods is export-controlled or not is the online system of Stanford University. However, this system requires users to have specialized knowledge of export control, which makes it difficult to use by researchers without extensive training. The goal of our research is to build a dialog system [Obayashi and Rzepka, 2019] in which experts in export control and artificial intelligence collaborate to develop a novel user-friendly support for non-experts. The system under development converts the text of export control laws and regulations into a computationally-processable format (an ontology), automatically makes inferences from articles to be judged, and is planned to add missing information through dialogue. Our system is to make it

possible for those who do not have knowledge of export control to easily and reliably determine whether technical information such as a potential publication is relevant or not. This should prevent the outflow of sensitive technology and thus contribute to national security. The system developed in this research will be provided free of charge and can be freely extended as open source.

1.2 Motivation for Sharing Data

As a small team for an ambitious project requiring combining various approaches, the authors hope that other researchers could use the data to test their algorithms and propose new methods. Nowadays new textual resources are prepared and widely shared but their character is concentrated on size, not quality, which is a natural consequence of current systems that require big data. Moreover, most of the widely shared datasets are in English causing that many NLP researchers from countries like Japan start working with English language instead of their native one.

2 Related Works

Plethora of textual datasets has been made available for both question answering and dialog processing. Recently open-domain question answering has gained popularity and many benchmarks were developed [Zhu *et al.*, 2021] and combining information retrieval with deep neural networks has become a popular research area [Abbasiantaeb and Momtazi, 2021]. However, most of the datasets originate in knowledge bases like Wikipedia and concentrate on factoids with simple answers. When it comes to queries requiring explanations, ranking existing replies is one popular method and Frequently Asked Questions have been often utilized for more than a decade [Wang *et al.*, 2009; Lin *et al.*, 2020] and the dataset we describe in this paper is similar to an FAQ. Another approach is to retrieve answers from linked data [Dimitrakis *et al.*, 2020], however legal documents like ours are just raw text and related methods cannot be used in a straightforward manner. If the data could be translated into graphs, there is a wide range of methods [Zheng *et al.*, 2018; Yasunaga *et al.*, 2021] but as semantic web research show, automatic ontology generation from text is not an easy task [Elnagar *et al.*, 2020].

When it comes to dealing with longer texts, such research has been comparatively less addressed, however recent works

*Contact Author / Equal Contribution

on long-form question answering has brought many new attempts to answer longer questions or generate longer answers when necessary. Still, the available datasets are dominated by short texts [Cambazoglu *et al.*, 2020] and even if there are dedicated ones for long-form question answering, they are also based on Wikipedia and provided mainly for English language [Kwiatkowski *et al.*, 2019]. For Japanese language a Wikipedia-based quiz question and answer data [Suzuki *et al.*, 2020]. Also a machine-translated SQUAD set [Rajpurkar *et al.*, 2016] has been made publicly available¹. [Takahashi *et al.*, 2019] have developed a QA dataset focusing on driving domain created from Japanese blogs. They constructed two wide-coverage datasets as a form of QA using crowdsourcing: predicate-argument structure one and a reading comprehension one. In the case of security export control, the domain is very narrow, therefore there is no easy access to the expert knowledge and traditional crowdsourcing is not possible.

3 Data Description

Our dataset is compiled from the guidance files prepared by the Center for Information on Security Trade Control (CIS-TEC)², and its target audience is those involved in security export control at companies, universities, etc., with the aim of providing them with the information they need to determine whether a shipment is relevant or not. Therefore, it is assumed that they have some basic knowledge of security export control and understand information on cargo in the field (e.g. chemical preparations). The FAQ data is provided within guidance pdf files³, therefore we manually retrieved questions and answers but omitted these with pictures and these referring to other documents without giving direct answers. In total we have extracted 548 question-answer pairs, which are available upon request. We have excluded pairs where images were used for asking a question or replying it. Some questions are long with short answers and vice versa, short questions require detailed replies. Due to the length constraints, here we analyze rather short examples (longer ones are given in the Appendix). The first one is about carrying a certain amount of a compound abroad:

Q: I am planning to take a very small amount of hydrogen fluoride (e.g., about 10g) to a foreign country. In this case, do I need to apply for an export license?

A: For substances that are used as raw materials for chemical preparations for the military, such as hydrogen fluoride, an export license application is required when taking them out of Japan and into foreign countries, even in very small quantities.

In the example above the question contains a direct specification of a question type (“do I need a license?”), for which yes/no answer can be determined (although an explanation

should be added) but it is not always the case. For example, let us consider the following Q&A:

Q: We are planning to export a storage container with teflon (tetrafluoroethylene resin) -coated wetted parts to China. In this case, how should we determine whether it is relevant or not?

A: Fluoropolymers listed in Article 2, Paragraph 2, Item 2 (c) of the Ministerial Order in Paragraph 3 (2) of Appended Table 1 of the Export Trade Control Order include tetrafluoroethylene, so those with a sealed structure are applicable.

In this case, the answer is labeled as “controlled” (meaning “requiring a license”), but this alone does not mean that it is strictly applicable. The reason is that the regulations also include a standard for capacity, for example, if the capacity does not exceed 0.1 cubic meter, it is not considered as “controlled”. This is due to the fact that the Q&A only takes into account important factors when making human judgments. An expert knows that fluoropolymers include tetrafluoroethylene resin, which is an effective piece of knowledge allowing human to focus on appropriate category. Similarly, such knowledge would be necessary when a machine performs a judgement, which makes this dataset a useful testbed for multi-hop question answering on long-form texts.

4 Data Annotation

All question-answer pairs are annotated a) by the first author who is a security export control expert and b) by a program matching numbers of articles and glossary terms included in the pairs.

4.1 Expert Annotation

We have decided to label the answers into four categories: “controlled”, “not controlled”, “requiring confirmation” and “other”. Example pairs for every of four categories is given in the Appendix. The expert has read all questions and related answers, then classified the type of answer with the labels above. There are 95 answers labelled as “Controlled”, 138 as “Not Controlled”, 139 as “Requiring Confirmation” and 176 as “Other”. Some of the answers were difficult to classify as for example the licence requirement was obvious but there was also a chance of some other problem occurring and the answer provided an advice as well. Such cases make labels “Requiring Confirmation” and “Other” simultaneously relevant. High number of “Other” suggests that granularity of the miscellaneous category could be increased further in the future. Many examples suggest that “defining”, “confirming” and “clarifying” could be used but the dataset is still not too large, therefore we decided to keep four labels.

4.2 Automatic Annotation

Japan Machinery Center for Trade and Investment⁴ publishes a 360 pages booklet containing trade security control terms linked to related article numbers. By courtesy of the Center we acquired an electronic version of this material and

¹<https://www.ai-shift.co.jp/techblog/1224>

²<https://www.cistec.or.jp/english/export/faq.html>

³https://www.cistec.or.jp/members/f_guidance/index.html
(registration is required to access the guidance files)

⁴<http://www.jmcti.org/jmchomepage/english/>

utilized the human-made terms for automatic keyword annotation. It must be noted that except specialized words as chemical compound names or names of viruses, everyday use words like “confirm” (*kakunin*) or “necessary” (*hitsuyō*) are also included. Except keywords we also have utilized article numbers from an XML file with related regulations⁵. It has to be noted that 20% of the question set contained only one or no keyword or article reference (0=8.2%, 1=10.9%)⁶. Because the average length of a question in the QA dataset is 110.16 ideograms (73.5 morphological tokens) and average number of nouns per question is 35.09, it can be again hypothesized that proper processing the data we prepared will require reaching beyond simple approaches.

5 Conclusions and Future Work

In this paper we have introduced a unique expert question answering dataset for the domain of security export control. Although relatively small (548 pairs), the developed data fill the gap in Japanese QA datasets by providing long-term questions and answers where both inquirer and answerer are experts with a different level of expertise. We have already performed set of experiments [Rzepka *et al.*, 2021] to see whether both classic (LDA [Blei *et al.*, 2003]) and neural (BERT [Devlin *et al.*, 2019]) approaches can correctly rank existing answers and find them in regulatory texts but the results show that they perform worse than a simple keyword matching. Because the set is small and the range of topics is very wide, fine-tuning is limited. However there are more sophisticated methods for both ranking, retrieving and generating answers to be tested. By providing this dataset we hope that the Japanese NLP community will use it for expanding current methods to deal with this unorthodox set of questions and answers annotated with answer intents, glossary words and article numbers.

6 Acknowledgements

This work was supported by JSPS KAKENHI Grant Number 20K12556.

7 Appendix

Examples of annotated questions and answers are given in Tables 1 and 2. Label abbreviations are: Con = “Controlled” (applicable), N-Con = “not controlled” (not applicable), “Conf” = “requiring confirmation”, “Other” = “Miscellaneous”.

⁵They are specified in the “Appended Table of the Foreign Exchange Order” (FEO) and the “Appended Table 1 of the Export Trade Control Order” (ETCO) provided by the Japanese government.

⁶One of the remaining problems is that article numbers can be written in formal and informal, full and partial manner, therefore full formal numbers we used cover only a small part of the data. We need to address this problem with manual approach or sophisticated regular expressions.

References

- [Abbasiantaeb and Momtazi, 2021] Zahra Abbasiantaeb and Saeedeh Momtazi. Text-based question answering from information retrieval and deep neural network perspectives: A survey. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, page e1412, 2021.
- [Blei *et al.*, 2003] David M. Blei, Andrew Y. Ng, and Michael I. Jordan. Latent dirichlet allocation. *J. Mach. Learn. Res.*, 3(null):993–1022, March 2003.
- [Cambazoglu *et al.*, 2020] B Barla Cambazoglu, Mark Sanderson, Falk Scholer, and Bruce Croft. A review of public datasets in question answering research. In *ACM SIGIR Forum*, 2020.
- [Devlin *et al.*, 2019] Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. BERT: Pre-training of deep bidirectional transformers for language understanding. In *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)*, pages 4171–4186, Minneapolis, Minnesota, June 2019. Association for Computational Linguistics.
- [Dimitrakis *et al.*, 2020] Eleftherios Dimitrakis, Konstantinos Sgontzos, and Yannis Tzitzikas. A survey on question answering systems over linked data and documents. *Journal of Intelligent Information Systems*, 55(2):233–259, 2020.
- [Elnagar *et al.*, 2020] Samaa Elnagar, Victoria Y. Yoon, and Manoj A. Thomas. An automatic ontology generation framework with an organizational perspective. In *53rd Hawaii International Conference on System Sciences, HICSS 2020, Maui, Hawaii, USA, January 7-10, 2020*, pages 1–10. ScholarSpace, 2020.
- [Kwiatkowski *et al.*, 2019] Tom Kwiatkowski, Jennimaria Palomaki, Olivia Redfield, Michael Collins, Ankur Parikh, Chris Alberti, Danielle Epstein, Illia Polosukhin, Matthew Kelcey, Jacob Devlin, Kenton Lee, Kristina N. Toutanova, Llion Jones, Ming-Wei Chang, Andrew Dai, Jakob Uszkoreit, Quoc Le, and Slav Petrov. Natural questions: a benchmark for question answering research. *Transactions of the Association of Computational Linguistics*, 2019.
- [Lin *et al.*, 2020] Jimmy Lin, Rodrigo Nogueira, and Andrew Yates. Pretrained transformers for text ranking: Bert and beyond. *arXiv preprint arXiv:2010.06467*, 2020.
- [Obayashi and Rzepka, 2019] Akihiko Obayashi and Rafal Rzepka. Towards interactive advisory system for security export control. In *Proceedings of IJCAI Workshop on Language Sense on Computer, Macau*, 2019.
- [Rajpurkar *et al.*, 2016] Pranav Rajpurkar, Jian Zhang, Konstantin Lopyrev, and Percy Liang. SQuAD: 100,000+ questions for machine comprehension of text. In *Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing*, pages 2383–2392, Austin, Texas, November 2016. Association for Computational Linguistics.

- [Rzepka *et al.*, 2021] Rafal Rzepka, Daiki Shirafuji, and Akihiko Obayashi. Limits and challenges of embedding-based question answering in export control expert system. In *Proceedings of the 25th International Conference on Knowledge-Based and Intelligent Information & Engineering Systems (to appear)*, Szczecin, Poland, September 2021. Springer.
- [Suzuki *et al.*, 2020] Masatoshi Suzuki, Jun Suzuki, Koji Matsuda, Kyosuke Nishida, and Naoya Inoue. Jaqket: Constructing a japanese question answering dataset from trivia questions (in japanese)aqket: Constructing a japanese question answering dataset from trivia questions (in japanese). In *In Proceedings of the 26th Annual Meeting of the Association for Natural Language Processing (NLP 2020)*, 2020.
- [Takahashi *et al.*, 2019] Norio Takahashi, Tomohide Shibata, Daisuke Kawahara, and Sadao Kurohashi. Machine comprehension improves domain-specific Japanese predicate-argument structure analysis. In *Proceedings of the 2nd Workshop on Machine Reading for Question Answering*, pages 98–104, Hong Kong, China, November 2019. Association for Computational Linguistics.
- [Wang *et al.*, 2009] Xin-Jing Wang, Xudong Tu, Dan Feng, and Lei Zhang. Ranking community answers by modeling question-answer relationships via analogical reasoning. In *Proceedings of the 32nd international ACM SIGIR conference on Research and development in information retrieval*, pages 179–186, 2009.
- [Yasunaga *et al.*, 2021] Michihiro Yasunaga, Hongyu Ren, Antoine Bosselut, Percy Liang, and Jure Leskovec. Qa-gnn: Reasoning with language models and knowledge graphs for question answering. In *North American Chapter of the Association for Computational Linguistics (NAACL)*, 2021.
- [Zheng *et al.*, 2018] Weiguo Zheng, Jeffrey Xu Yu, Lei Zou, and Hong Cheng. Question answering over knowledge graphs: question understanding via template decomposition. *Proceedings of the VLDB Endowment*, 11(11):1373–1386, 2018.
- [Zhu *et al.*, 2021] Fengbin Zhu, Wenqiang Lei, C. Wang, Jianming Zheng, Soujanya Poria, and Tat-Seng Chua. Retrieving and reading: A comprehensive survey on open-domain question answering. *ArXiv*, abs/2101.00774, 2021.

Question	Answer	Label
In order to eliminate multipath signals from radio and TV broadcasts, etc., a receiver (frequency range 1.5 MHz to 87.5 MHz) has been designed using adaptive interference signal suppression technology. The design is capable of suppressing multipath signals beyond 15 decibels. Is this design technology regulated by the technology specified in Article 21, Paragraph 2, Item 3-2(d)(3) of the Ministerial Ordinance?	Article 21, paragraph (2), item (iii)-(d)(3) of the Ministerial Ordinance regulates the technology required for the design or manufacture of radio transmitters and radio receivers (frequency range 1.5 MHz to 87.5 MHz) designed to suppress disturbance signals beyond 15 decibels using adaptive interference signal suppression technology. (2) A multipath signal is not an interference signal because it is a signal that is reflected or diffracted from the desired signal by obstacles such as buildings and terrain. Therefore, if a cargo design technology uses adaptive interference signal suppression technology that is effective only for multipath signals, it is “not applicable” to Article 21, Paragraph 2, Item 3-2(d)(3) of the Ministerial Ordinance. However, if the adaptive interference signal suppression technology is equally effective against disturbance signals and the cargo is capable of suppressing disturbance signals in excess of 15 decibels, the technology required for its design or manufacture is “applicable”.	Con
We are planning to set up a subsidiary in a foreign country to manufacture cross-flow filtration systems and components (membrane modules, etc.) in order to be cost competitive. This subsidiary will be operated solely with the capital of a Japanese company, but do we need a permit?	Cross-flow filtration devices and components (e.g., membrane modules) are list-controlled goods, and their design, manufacture, and use technology are subject to regulation. In addition, even joint ventures and proprietary subsidiaries are considered non-residents, and the provision of technology to non-residents is subject to regulation. The residency of a corporation, etc. shall be determined based on whether or not it has its principal office in Japan, and the residency of a branch office, sub-branch office or other office of a corporation, etc. shall be determined as follows: (1) Branches, branch offices and other offices of Japanese corporations, etc. located in foreign countries shall be treated as nonresidents. From “Interpretation and Application of Foreign Exchange Laws and Regulations”.	Con
We have provided a numerical control program built into a machine tool that falls under the list regulations using the special exception in Article 9, Paragraph 2, Item 14 (c) of the Ministerial Ordinance on Foreign Trade. However, we found out that the program has a bug. We would like to provide a corrected program, but since it is not built into the cargo this time, we cannot use this special exception. In this case, do we need to obtain a new service transaction permit?	According to the results of public comments at the time of the revision of the Ministerial Ordinance in August 2012, the view was expressed as “Programs provided with a service transaction license” in Article 9, Paragraph 2, Item 14 (d) of the Foreign Trade Ordinance includes programs provided by applying the special provisions that do not require a license. Since this case falls under this category, the special exception in Article 9, Paragraph 2, Item 14 (d) of the Foreign Trade Ordinance can be applied and no new service transaction license is required.	N-Con
One of our products is a lubricant oil. This product is mainly composed of ingredients regulated in Section 5 (12), but it is designed and manufactured only for use as a lubricant, and we have not confirmed its actual use as a refrigerant in electronic equipment. Is it correct to assume that such products cannot be used as refrigerants, and that they do not fall under item 5(12)?	If it is judged that the product cannot be used for the regulated application based on the specifications at the time of manufacture and actual examples of use, it may be judged as not applicable. We recommend that you keep the information that forms the basis for your judgment in some form, such as by storing it in writing together with the documents used to determine whether the product is applicable or not.	N-Con

Table 1: Examples of annotated questions and their answers for categories “Controlled” (Con), and “Not Controlled” (N-Con)

Question	Answer	Label
What should I do if I don't know how much of the relevant chemical substance is contained in the cargo I intend to export?	In the case of a trading company, the information should be based on the information exchanged in normal business practices (MSDS, catalog, etc.). If the chemical substance does not fall under any of items 1(3), (13), 2(3), or 4(6), and does not exceed 10% of the price of other goods in the mixed destination, it is not considered to be a relevant chemical substance. Normally, the price of a small amount of a substance that is not listed on the MSDS, etc. is not considered to exceed 10% of the total price of the shipment. However, if there is little information on the chemical substances contained in the product and you do not know what the main substances are, it is necessary to confirm the applicability by contacting the manufacturer.	Conf
How should diaphragm type vacuum pumps and bellows-type vacuum pumps be judged as to whether they are applicable or not? Are they classified as "vacuum pumps" or "seal-less pumps"?	Diaphragm type vacuum pumps and bellows type vacuum pumps can usually only handle gas and cannot handle liquid. To that extent, none of the pumps fall into the category of normal seal-less pumps that can handle liquids. Determine the applicability of the pump as a vacuum pump. However, if the pump can also handle liquid, it must also be judged as a seal-less pump. (Canned pumps, magnet pumps, bellows pumps, and diaphragm pumps are classified as "seal-less pumps".)	Conf
Does the provision of "single or composite oxides of zirconium and composite oxides of silicon or aluminum" in Paragraph 5(3) of the Appended Table of the Foreign Exchange and Foreign Trade Ordinance and Article 17, Paragraph 3, Item 1(a)(1) of the Ministerial Ordinance on Freight, etc. regulate four types of oxides: 1) single oxide of zirconium, 2) composite oxide of zirconium, 3) silicon, and 4) composite oxide of aluminum?	No, it doesn't. The same regulation regulates four types of oxides: 1) single oxides of zirconium, 2) complex oxides of zirconium, 3) complex oxides of silicon, and 4) complex oxides of aluminum. The English text of the Wassenaar Arrangement states "Single or complex oxides of zirconium and complex oxides of silicon or aluminium".	Other
If a model for which the declared value was accepted before November 19, 2009, and has not been produced for a considerable period of time in the past, how should the declared value be handled?	In accordance with the provisions of 12) Others (3) of "Declared Values for Linear Axis Positioning Accuracy" (Export Notes 21, No. 49, 2009/11/13 Trade Bureau No. 3, dated September 27, 2013), please submit the following documents. Documents proving that production has been discontinued (if the possibility of resuming production in the future cannot be denied, please include the following: a copy of the checklist acceptance sheet (1 copy) and the original of the previous declaration value acceptance sheet.	Other

Table 2: Examples of annotated questions and their answers for categories "Requiring Confirmation" (Conf) and "Miscellaneous" (Other)