

NorOpenBookQA and NorCommonSenseQA guidelines

While we begin with the OpenBookQA dataset, these guidelines will be applicable for both this and the CommonsenseQA dataset. We begin with a short introduction of the datasets, before discussing some common problems.

Datasets

We will first briefly introduce the two datasets for Question Answering (QA).

OpenBookQA

The OpenBookQA dataset (Mihaylov et al., 2018) is a Question Answering (QA) dataset. It differs from the other dataset that we will be looking at, CommonsenseQA, in that for every question, a fact is provided representing important knowledge that is needed to answer the question (in addition to natural language understanding and common sense reasoning).

Which of these would let the most heat travel through?

- A) a new pair of jeans.
- B) a steel spoon in a cafeteria.
- C) a cotton candy at a store.
- D) a calvin klein cotton hat

Fact: Metal is a thermal conductor

CommonsenseQA

The CommonsenseQA dataset (Talmor et al., 2019) contains questions created in a way so that "common sense" is required to answer questions. Specifically, it is made in such a way that the answers cannot be inferred from the question, and that world knowledge at the common sense level is required to solve it.

I'm crossing the river, my feet are wet but my body is dry, where am I?

✗ waterfall, ✗ bridge, ✓ valley, ✗ bank, ✗ island

Format

You will be provided with xlsx-files that contain data looking something like this, but with varying number of alternatives, or with or without facts:

Source ----- Alternative A ---- Alternative B ---- Alternative C ---- Fact ---- Label

Your translations and localizations come directly under the original English. You can add several localizations if you feel that there are several possible options. After the last English sentence in your document, you can add inspired sentences. Make sure to include a blank row before these sentences.

Your final document can look something like this:

Source ----- Alternative A ---- Alternative B ---- Alternative C ---- Fact ----

Localization ...

Translation ...

Localization ...

Source ----- Alternative A ---- Alternative B ---- Alternative C ---- Fact ----

Source ----- Alternative A ---- Alternative B ---- Alternative C ---- Fact ---- U

Source ----- Alternative A ---- Alternative B ---- Alternative C ---- Fact ----

Translation ...

Source ----- Alternative A ---- Alternative B ---- Alternative C ---- Fact ----

Localization ...

Inspired question from sentence 1

Inspired question from sentence 3

Inspired question from sentence 3

Inspired question from sentence 4

The first sentence has three sentences: Two localizations and one translation. The second sentence is marked as U, and is unused, both for translation, localization and inspiration. The third sentence is an inspired sentence, but no direct translations or localizations. The fourth sentence has a single translation. The fifth sentence has a single localization.

In the training round, the "label"-column was unlabeled, but it will be clear when you get new projects.

Annotation procedure

To annotate, you have three options. You can translate, with or without localization, you can leave a sentence unused due to difficulties, errors or problems (labelled as U), or you can be inspired by sentence to write a new question, in which case you do not write anything below the sentence itself, but add them to the end of the document.

If you are comfortable with translating, or translating and localizing a sentence, you are free to do so. Some of you might be tempted to add additional comments or clarifications. You may do that on separate files, but please make sure that there are no additional fields in the files you hand in, as these will be automatically treated later on. If you have comments, whether they are things you are

uncertain about, or things you think we should know about, you can create an issue, or send us an e-mail.

Translation and localization

Translation is no easy task. There are many ways to translate a sentence or a set of sentences, and many things to keep in mind. Feel free to use the Issues in this Github repository to ask questions. It is better to ask, and to discuss potential problems, than to ponder alone for a long time.

Tools

You are encouraged to use various dictionaries, such as ordbokene.no, and other tools such as naob.no, snl.no or even wikipedia. However, we want to stress that you cannot use other machine translation tools. The point of this annotation effort is to provide more natural translations, and it is better to introduce some human-like errors, than using machine translation. It is exactly because machine translation works so poorly that we want to focus on this task.

Style

We will translate into both Bokmål and Nynorsk (depending on your preferred variant, as specified in the submitted form). When it comes to further variation, you should stick to a norm that is not too stiff and alienating (i.e. close to how you would normally express yourself in written form), but also not too radical. Think jenta, boka and hytta, but selv, skog and vei, not sjøl, skau and veg in Bokmål. However, you are fairly free otherwise. For example, if you really do not use the feminine when writing, there is no need for you to look it up. Postpositioned possessors are preferred (boka mi, bilen min, vannet mitt), but this is not so strict. The -et ending is preferred in the preterite and perfect (snakket, not snakka).

Wherever possible, choose the more Norwegian alternative. Automatically translated text often has plenty of anglicisms (f.eks. komme opp med fra "come up with" i stedet for finne på). You should choose a Norwegian word wherever possible, especially if that is the term used in a Norwegian context, or if it can be argued that more people would intuitively understand what something means. For example, use lysbrytning instead of refraksjon, but you do not need to write fjernsyn for TV. Use your common sense 😊.

You should always follow the norm. Refer to the website of [the Norwegian Language Council](#) and [Ordbøkene](#). Note especially the section on comma rules. Note that NAOB is not a normative dictionary, but it is still great, so if you cannot find something in Ordbøkene, check there. If you are uncertain, use Github Issues or ask us. When it comes to scientific vocabulary, it is better to check which terms are in actual use, if you are uncertain. Store Norske Leksikon (SNL) is a great source of accessible scientific vocabulary.

Titles and names

If a show or movie is mentioned, you can use the Norwegian name if it exists and is commonly understood. This means that the Lion King would be "Løvenes konge", but Die Hard does not have to be "Operasjon Skyskraper". Titles of songs are not translated.

Relation between questions and possible answers

Make sure that you read the questions together with the answers you translate, to make sure that they sound OK. It is easy to forget to add correct grammatical markers, or perhaps add subjects where the English sentences did not have them.

Read sentences aloud

A tip is to read the sentences you write aloud, especially together with the replies. This can make it easier to hear if there are any inconsistencies.

Examples

Translation and localization

If there are cases where a sentence mentions something that is generally not part of Norwegians' knowledge, or something that is very specific to the US. For example, sentences asking about where a certain place in the US is located, or a fact containing culture-specific background knowledge. In these cases, you can rewrite the sentence to fit a Norwegian context. A localized sentence is typically a translated sentence (or set of answers), but where one specific item or context is changed, typically due to its specific cultural connotations. Common examples include names of unknown biological entities, place names, local geography, certain products, celebrities, and many more.

In the example below, one of the possible answers is Kool-aid, a popular drink mix powder from the US. While Norwegians might know about Kool-aid through interaction with american culture, it might still be unknown for many. We might want to rewrite it using a word that has the same function. It is possible to do changes both in the questions and the answers, but not in the facts. The logical relationship between the question, answers and fact must be preserved.

Which is a solid? A) Kool-aid B) urine C) helium D) ice sculpture
Fact: Matter in the solid phase has definite shape

A possible Norwegian rendition could be as follows. We use the term "saft" instead of "Kool-aid".

Hvilken av disse er i fast form? A) Saft B) Urin C) Helium D) En isskulptur

In some cases you might have to make larger changes. You are free to provide both a localized and a translated version. This is especially true if, even if the English sentence is comprehensible in a Nordic context, there might be even more suited candidates. For example, in the initial training round, a question involved a piranha (piraja). One can assume that most Norwegians know about this fish, and that it lives in rivers, which is the only fact about the fish necessary to answer the question. But since any water-dwelling animal would have kept the feeling of the original English question, it could easily be substituted, as some of you did, by more Norwegian water-dwellers: ørret, gjedde, etc.

Unused (Erroneous, difficult, ...)

In some cases, the English sentence is simply difficult to understand, nonsense or of a type that is very difficult to translate naturally. In these cases, you do not need to dwell on them, you can simply give the sentence the label E for Error. We would of course like you to not skip sentences lightly, but at the same time there is no benefit from having very awkward or unnatural Norwegian sentences, if this can be avoided. You can also skip them if you feel that they are just difficult to translate, or if they would require you to check many things before translating. Remember that sentences in this case can also be used for inspiration, described in the next paragraph.

Your own examples

Sentences that would otherwise be marked as U, but that trigger the imagination in some fashion, may be used to create own sentences. For example, say you are given the following original sentence:

Webbed feet are used to move faster through what? A) rain B) winds C) clouds
D) swamps webbed feet are used for moving faster through water by aquatic animals

But you cannot think of what webbed feet (svømmeføtter) are, or maybe you find the alternatives strange. But you see some of the words, and it causes you to create the following questions: (The letters (A, B, etc.) are added for readability, you do not add them in the document)

Hvilke av disse dyrene har svømmeføtter? A) Ender B) Hunder C) Løver D) Hester

Hvis jeg går over ei myr, er det mest sannsynlig at skoene mine blir A) spist opp av kjøttetende planter i myra B) våte C) tørrere enn de var D) magiske

Hvilke av disse dyrene svømmer raskest? A) Voksne blåskjell B) Tiur C) Muldvarp D) Laks

There is no fixed limit to the number of creative questions you can write, but we ask you not to make too many where only one or a few words are changed, to promote variety. The new sentences are not marked specifically. You are free to, and encouraged to create inspired sentences from English sentences that you have already translated and/or localised.

Specific types of errors

Finally we will look at some types of errors, and some tips about how to deal with them.

Errors or problems with the original English sentences

In some cases, there are problems with the original sentence that make the translation work difficult. For example, there may be grammatical mistakes, problematic formulations etc. If the issues are serious enough, you can discard the sentence (mark it with U) as discussed above, but if not, here are some tips.

We do not change the original sentence even if we find something wrong with it. The English sentences are always left as they are.

Spelling mistakes

If the English contains a spelling mistake, but the meaning is clear, you do not need to transfer the spelling mistake to Norwegian. Translate or correct the sentence as if it had no mistakes. This type of mistake does not need to be marked.

In the example below, the word "to" is spelled "tho". This is left in the original sentence, but the Norwegian sentence should not contain any similar spelling mistakes.

Magma piled on the Pacific bottom tho form}

Linguistically guided answers

Some of the English questions are answerable solely based on the word class or form of the answers. In these cases, try and see if it is possible to restructure the question so it either 1) works for Norwegian, or 2) does not have the same structure. This might also happen in Norwegian, where for example the question ends in an article that identifies the gender of the correct answer word.

If your opponents in a marathon are very fast it is easy to fall what? (A) lightly (B) behind (C) train (D) slowly (E) sitfast

In the sentence above, we only need to know that "fall behind" is a set phrase. If these become too difficult, you can also discard them.

Factual underspecifications

Some questions are not specific enough, and contain more than one correct answer. In the example below, the question is about pasteurization, perhaps based on the wrong notion that this process is only used for dairy. The provided correct answer is B, which is the most obvious, but pasteurization would be able to reduce the amount of bacteria of all the products that are possible replies, and in fact is commonly used for juice. While these can be frustrating, try and translate them as they are, or create an inspired question without the ambiguity.

Pasteurization reduces the amount of bacteria in (A) green tea (B) cheddar cheese (C) orange juice (D) almond milk

Stereotypes and bias

Some questions contain premises that lead to unfortunate conclusions, where certain types of biases and stereotypes appear. For example, the following question has (C) as its correct answer, having the underlying common sense assumption that Mexican food leads to digestive issues. Questions relying on negative stereotypes like this should be labeled 'stereo'.

Billy was a human, and he had human bodily functions. He needed to catch his flight but ended up missing it, because he had too much Mexican and was stuck where? (A) bed (B) couch (C) bathroom (D) school (E) airport

Sources

1. [CommonsenseQA: A Question Answering Challenge Targeting Commonsense Knowledge](#) (Talmor et al., NAACL 2019)
2. [Can a Suit of Armor Conduct Electricity? A New Dataset for Open Book Question Answering](#) (Mihaylov et al., EMNLP 2018)