Consistency Annotation Guidelines

- You are given pairs of English/German sentences. Both are generated automatically, so both sides may contain errors. There is no directionality between both sentences, both are considered as translations of each other.

- Your task: assign a consistency score to each sentence pair.

  - **Definition**: Perfect consistency between both sentences is achieved if both are **semantically equivalent**, with a preference for a **faithful** translation approach, meaning that **stylistic, lexical, and grammatical characteristics** should be transferred whenever possible within the grammatical constraints of each language.

  - Importantly, a sentence pair can be ungrammatical, but still be considered perfectly consistent. In such a case, consistency means that both sentences match on word or phrase level, even though on neither side do the words/phrases add up to a grammatical sentence (that is: similar grammatical errors on both sides should not hurt consistency).

- To evaluate consistency, choose one out of 4 scores:

  - 4 "OK": Both sentences mean virtually the same thing, and follow a similar style and sentence structure. Includes ungrammatical cases that have similar issues on both sides. Minor deviations in nuances are ok, and minor grammatical issues are ok even if occurring only on one side.

  - 3 "close": Both sentences mean virtually the same thing, but there are certain inconsistencies with regards to **style, spelling, or sentence structure** that should be improved upon.

  - 2 "problematic": The **meaning** of sentences overlaps only partly, while other parts do not mean the same thing. Typical cases are when e.g. the main clause is consistent but a relative clause has a different meaning; when only one sentence has major grammatical issues; when both sentences are ungrammatical and these issues are considerably different on both sides.

  - 1 "catastrophic": Both sentences have only little in common.

- Case of **proper names** (people, places, organizations, ...): Spelling variations between German and English sides should be counted as minor stylistic issues, not as major semantic issues. To be precise: By spelling variation we mean a change to the spelling that has only negligible impact on the pronunciation. Such inconsistencies should be given a score of "3".

Typical examples

Note that there may be cases that fall in between categories and require a judgment call

"OK" examples.

score	English	German	comment
4 "OK"	OK, a fond that doesn't do its fish, a phone, the measures of success, by the success of its predators.	Okay, ein Spaß, der seinen Fisch nicht tut, ein Telefon, die Maßnahmen des Erfolgs seiner Raubtiere.	*major grammatical issues, but consistent on both sides*
	A fish farmed but also a bird sanctuary.	Ein Fischfang, aber auch Vogelschutzgebiet.	*"fish farmed" is odd, but does match the meaning of the German side*
	(Audio) NR: Yeah.	(Audio) NR: Ja.	*Sentence in video-subtitle style, "NR" might be an abbreviated name. Regardless of perhaps unclear meaning, both sides are perfectly consistent.*
	Almost never cook with it.	Fast nie koche damit.	*Example for a minor disfluency that should be forgiven.*
	A crucial part of this case study was that a teachers were reduced too.	Ein wichtiger Teil dieser Studie war, dass auch Lehrer auch reduziert wurden.	*some grammatical issues, but similar on both sides*

"Close" examples.

score	English	German	comment
3 "close"	Above that —	Darüber hinaus.	*Small, stylistic diverge in choice of punctuation*
	Also, these are room mates and sources. The son doesn't China all day, doesn't China every day, and likebytes, the wind doesn't go all the time.	Das sind auch Mitbewohner und <unk>uellen. Der Sohn macht nicht jeden Tag China, nicht jeden Tag, und wieschmettert der Wind nicht die ganze Zeit.	*While odd, both match semantically. The only issue is the "<unk>" on the German side, which does not destroy understandability, but should be improved upon.*
	Doctor King made that shift in thinking.	Dr. Keeng machte diese Verschiebung in der Denkweise.	*Spelling variants and inconsistent abbreviations would count as minor stylistic problems.*
	Five minutes ago.	Vor fünf Minuten zuvor... Fünf Minuten her.	*The repetition on the German side would count as a stylistic mismatch.*

"Bad" examples

score	English	German	comment
2 "bad"	Actually, save the entire civilization.	Die gesamte Zivilisation rettet tatsächlich.	*The words match, but both do not convey the same meaning, as the English side is grammatical but the German side is ungrammatical and difficult to understand.*
	Now, what about this?	Nun, wie geht 's?	*While there are similarities, considerable portions of the sentences have different meanings.*
	Finally, the game we're launching at March 3rd.	Schließlich, das Spiel dass wir im März Drittel starten.	*While there's significant overlap, the temporal expressions have different meanings on both sides.*
	First year, this property had 600,000 birds on it.	Im letzten Jahr hatte er 600.000 Vogelarten, mehr als 2100 000 unterschiedliche Dinge.	*The German side has more information than the English side.*

"Catastrophic" examples

score	English	German	comment
1 "catastrophic"	Path theories. That was 16 years ago. They played Diskmes for 18 years.	Das war vor 60 Jahren. Diese Polizei bestand aus 18 Jahren.	It's hard to find any overlap in meaning.
	So CEOs, a little bit better than average, but here's where it gets interesting.	Selbstverständlich verstärkt sich die Strukturen.	It's hard to find any overlap in meaning.
	Now, what about 1,000 dollars?	Nun, ich möchte, dass Sie das tun.	This is a bit of a corner case as the first word is shared, but other than that it's hard to find any overlap.
	Pay gamers and embryos and electronics.	Papierspieler sind Virtuosen und Entwicklungsländer.	It's hard to find any overlap in meaning, aside from "gamers".