

# Translationese Detector

How to find errors in translated articles

Andrea, Alexandra, Dominick, Manuel, Philip, Sharon, Susan

**#SWIHACK**

# What is Translationese? 🤖

Im Roman geht es  
prinzipiell um einen  
Schwarzarbeiter, dessen  
Schmuck und Handy gestohlen  
wurde. Der Dieb war aber  
nicht sehr genial, denn er  
setzte die Küche in Brand  
und blamierte sich.

The roman is in principal  
about a blackworker, whose  
schmuck and handy was  
stolen. However, the thief  
did not act genially, he  
set the kitchen to brand  
and blamed himself.

# Also this is Translationese...

**Geheimdienst-Affäre: Mehrere Staaten  
gelangen in Crypto-Skandal an die  
Schweiz**

Nach Enthüllungen über das  
Ausspionieren Dutzender Länder durch  
Geheimdienste aus Deutschland und den  
USA mittels manipulierter  
Verschlüsselungsgeräte (...) haben  
mehrere Staaten bei der Schweiz  
interveniert. Dies sagte ein Sprecher  
des Schweizer Verteidigungs-  
departements (VBS) zu Tamedia.

Unklar bleibt, welche Länder  
intervenierten, und wie genau die  
Reaktionen ausfielen.

**Secret service affair: Several  
states get to Switzerland in  
crypto-scandal**

After revelations about the spying  
of dozens of countries by secret  
services from Germany and the USA  
using manipulated encryption  
devices (...), several states  
intervened with Switzerland. A  
spokesman for the Swiss Defence  
Department (VBS) told the Tamedia  
newspapers this.

It therefore remains unclear which  
countries intervened and how  
exactly the reactions were.

# Translations can go wrong in many ways

Expressions and word play – Is there an equivalent?

Quotes – Do people really talk like that?

Adjectives – Is the translated adjective or adverb too weak or too strong?

# Translations can go wrong in many ways

Tenses – Did the tenses get translated correctly?

Word count – Is the text much longer/shorter than the original?

Is it a case of “false friends” – similar look, but different meaning?

# What are false friends?

sensibel ≠ sensible

Perspektive ≠ perspective

aktuell ≠ actual

Marke ≠ mark

Rock ≠ rock

Schmuck ≠ schmuck

# What we did: Focus on word count per sentence & false friends

## Word count

- ❖ English texts are usually shorter than German texts
- ❖ If the English text is longer than the German original: flag
- ❖ measurement:

difference >4: warning

difference >10: error

## False Friends

- ❖ Based on a dictionary
- ❖ Warning message according to the number of false friends per sentence

# The Translationese Detector