**Korean War**

Korean War, conflict between the Democratic People's Republic of Korea (North Korea) and the Republic of Korea (South Korea) in which at least 2.5 million personnel lost their lives. The war mark- ed international proportions in late 1950 when North Korea, supplied and defended by the Soviet Union, invaded the South. The United Nations, with the United States as the principal participant, joined the war on the side of the South Koreans, and the People's Republic of China came to North Korea's aid. After more than a million combat casualties had been suffered on both sides, the fighting ended in July 1953 with Korea still divided into two states.

**Subject Data**

115 items of handwriting annotations on the cover of the photographs: 4 items of miniatures (unrecorded magazines, torn BC). Created digital image files of them by scanning.

**OCR (Optical character recognition)**

Erased Illustrations of the magazines, Level the angle, and Adjust the brightness. 4 miniatures divided into 96 items.

**Preprocessing**

Separated vertical columns and borders.

**Art – OCR: Google Cloud Document AI**

This is a document processing task that requires an OCR, Processor and Handwriting Recognition in Turkish. It gives three files that covers us to run the program. In return, we transformed this structured information into text data in CSV.

**Introduction**

Harp (Korea and War) (1954)

In return, we transformed this unstructured data to check which we hired a professional translator language resources that can be used for evaluation. In preprocessing steps, BLD can measure the similarity between a generated paraphrase and the original sentence.

**Process**

Preprocessing before OCR

- Line breaks and tables require an extra preprocessing procedure.
- Poor print quality led to poor recognition.
- Handwriting recognition was also poor.
- Failed to recognize documental marks.
- Time consuming post-processing.

Reviewing the OCR output corresponds to the preprocessing work for the translation. We refined the outcome within the above-mentioned limitations and generated source text files for machine translation.

**Preprocessing Machine Translation**

**Evaluating Deep and Google Translator**

**Solution**

We decided to create a process using AI. We also expected the effects of time saved and cost reduction.

**Challenge**

The collection at HMM library has recently acquired the personal archive of Major Sjakke Dobbe, a Turkish military officer who served in the Korean War. The archive includes the following materials:

1. Harp (Korea and War) (1954)
2. Turkey/Turkey’s participation in the Korean War
   At the United Nations Security Council held at the request of the United States, North Korea’s military actions were defined as aggression, and the United Nations therefore requested North Korea to stop its military actions and withdraw from South Korea. As North Korea refused, the UN decided to send troops to the Korean War. In accordance with this decision, the Turkish army entered the war as part of the UN forces and was one of the four countries that sent a large-scale brigade-level army. Turkey was the second country to decide to enter the war, following the United States. During the Korean War, Turkey dispatched 21,212 people and in the end, 2,322 people were killed, 1,155 people were injured, and 244 persons lost their lives. The war reached international proportions in June 1950 when North Korea, supplied and defended by the Soviet Union, invaded the South. The United Nations, with the United States as the principal participant, joined the war on the side of the South Koreans, and the People’s Republic of China came to North Korea’s aid. After more than a million combat casualties had been suffered on both sides, the fighting ended in July 1953 with Korea still divided into two states. Negotiations in 1954 produced no further agreement, and the third line has been accepted ever since as the de facto boundary between North and South Korea.

**Machine Translation**

The IG0N file obtained from OCR text extraction went through a preprocessing process for machine translation. While using OCR, there were a few cases where broken sentences happened in the middle of a sentence. In those cases, machine translation was not natural, so we took caution and went through the preprocessing process.

**Conclusion**

Because the original text was in Turkish, it was impossible to confirm whether it had been translated accurately. So, we decided to use BLEU score which can evaluate machine translation.

**Evaluating**

In addition, after a survey of native English Speakers. To four evaluation criteria: whether it was fully translated into English, natural, appropriate words, and grammatical errors. Then we asked “Why do you choose this option?”

On the other hand, if you were interested in getting a full translation of a source text, you might consider using a commercial translation service or a professional translator. A professional translator can provide a more accurate and natural translation, taking into account cultural and linguistic nuances. They can also provide additional notes and suggestions. Once you have the translated text, you can use various tools to further refine and improve the translation. This might include checking for errors, making adjustments to grammar and syntax, and ensuring the translation sounds natural and appropriate. Professional translators often use tools like translation memory systems or terminology databases to ensure consistency and accuracy in their work.