



Improving Readability: Evaluating and Applying Sequence-to-Sequence Transformer Models for Simplifying Dutch Governmental Letters

Master Thesis

Applied Data Science

Department of Information and Computing Science

First examiner:

Dr. Florian Kunneman

Second examiner:

Dr. Pierre Albert

Candidate:

Ole (Willem) Mathijssen
6576281

In cooperation with:

Rijksdienst voor Ondernemend
Nederland (RVO)

July 7, 2024

Abstract

Text simplification aims to make text more readable by reducing the linguistic complexity. This study explores the use of sequence-to-sequence transformer models to simplify Dutch governmental letters, enhancing their readability for individuals with low literacy levels. Various models, including T5-Small, BART-Base, mT5-Small, mBART-Large-50, T5-Base-Dutch, UL2-Small-Dutch and UL2-Small-Dutch-Simplification, were trained on datasets comprising complex and simplified Dutch sentences. These models were evaluated using quantitative metrics such as the Flesch-Kincaid Grade Level, BLEU score and SARI score, complemented by a qualitative analysis. The best-performing model was applied to a dataset of letters provided by the Rijksdienst voor Ondernemend Nederland to produce simplified versions. The study demonstrates that while the models slightly improve readability as indicated by Flesch-Kincaid scores, qualitative analysis reveals significant issues with content preservation and coherence. This highlights the need for further refinement to achieve the desired readability improvement, while maintaining accuracy in Dutch text simplification.

Contents

1	Introduction	1
1.1	Context	1
1.2	RVO and Research Objective	1
2	Literature Review	2
2.1	Related Work	2
2.2	Theoretical Framework	3
2.2.1	Sequence-to-Sequence Model Architecture	3
2.2.2	Transformers	3
3	RVO Dataset	4
3.1	Data Description	4
3.2	Data Cleaning	5
4	Methodology	6
4.1	Training and Evaluation Datasets	6
4.1.1	Municipal Dataset	6
4.1.2	UWV Dataset	7
4.2	Model Selection and Training	7
4.2.1	Models Selected	7
4.2.2	Training Approaches	8
4.2.3	Hyperparameter Tuning	8
4.3	Sentence Simplification	9
4.4	Evaluation Metrics	9
4.4.1	FKGL (Flesch-Kincaid Grade Level)	9
4.4.2	BLEU Score	10
4.4.3	SARI Score	10
4.5	Rewriting RVO Letters	10
4.5.1	Segmentation and Simplification of Letters	10
4.5.2	Evaluation of Simplified Letters	10
4.6	Implementation	11
5	Results	11
5.1	Performance of Models on Test Set	11
5.2	Comparative Analysis	12
5.2.1	Quantitative Analysis	12
5.2.2	Qualitative Analysis	12
5.3	Model Selection for RVO Simplification	14
5.4	Evaluation of Simplified RVO Letters	14
5.4.1	Quantitative: FKGL Scores	14
5.4.2	Qualitative: Inspecting Rewritten Letters	16
6	Discussion	17
6.1	Interpretation of the Results	17
6.2	Implications for Readability Improvement	18
6.3	Limitations of the Study	18
6.4	Suggestions for Future Research	19
7	Conclusion	19
	References	22

A	Appendix	23
A.1	Letters with Highest FKGL Scores	23
A.2	Letters with Average FKGL Scores	27
A.3	Letters with Biggest FKGL Score Improvements	31
A.4	Additional Tables and Figures	34

1 Introduction

1.1 Context

Text simplification (TS) is a field within natural language processing (NLP) focused on making written content more accessible. By reducing the linguistic complexity of a text, TS aims to make it easier to read and understand while maintaining the original content and meaning (Alva-Manchego et al., 2020; Shardlow, 2014). Text simplification is useful, as it can benefit individuals with low literacy levels, people who lack reading comprehension skills and non-native speakers by making written content easier to understand (Al-Thanyyan & Azmi, 2021).

Sentence simplification, closely related to text simplification, focuses on the modification of individual sentences rather than entire paragraphs or texts. This involves transforming complex sentences into simpler ones by breaking down long sentences, replacing complex words with simpler alternatives and rephrasing sentences to enhance readability (Seidl & Vandeghinste, 2024). Research specifically addressing text simplification at a text level is limited, with most studies focusing on sentence simplification (Alva-Manchego et al., 2020; Martin et al., 2022). It should be clarified that most works which mention text simplification actually refer to sentence simplification, and in the context of this research the terms are used interchangeably (Štajner, 2021).

Text simplification is of particular importance in the Netherlands because many individuals face difficulties with literacy. Over 2.5 million out of 14 million people over the age of 16 are considered low-literate, meaning they struggle with reading or writing (Hobo et al., 2023). A low literacy rate poses a serious barrier to accessing information and can be especially problematic when people do not understand official communications from government agencies. Previous research has shown that texts written by government agencies are not well understood by a significant portion of the Dutch population. And despite various initiatives, which include language programs and return stickers for letters that are unclear, the need for an effective solution remains (Harmsen & Raaij, 2023).

Sequence-to-sequence (seq2seq) models are widely used to address various text-to-text generation problems, including text simplification (Alkaldi, 2022; Alshantiri et al., 2023). These models transform a sequence of words into a simpler sequence to enhance readability, while trying to maintain the original meaning. The use of pre-trained seq2seq models, have shown promise in effectively simplifying Dutch sentences (Seidl & Vandeghinste, 2024).

1.2 RVO and Research Objective

The Rijksdienst voor Ondernemend Nederland (RVO) is a government agency in the Netherlands responsible for supporting entrepreneurs and organizations (Rijksdienst voor Ondernemend Nederland, 2024). To enhance the accessibility and readability of their official letters, RVO has provided a dataset of 131 letters that need simplification. The aim is to simplify these letters automatically, so that individuals with low literacy levels can better understand the information they contain.

The focus of this research is on utilizing automatic sentence simplification in Dutch, to simplify text. More specifically, this work aims to clarify the following question.

Research Question: *To what extent can sequence-to-sequence transformer models improve the readability of Dutch official letters provided by the Rijksdienst voor Ondernemend Nederland?*

To answer this question, two Dutch datasets are utilized, both containing complex sentences and their simplified counterparts. One dataset contains communication from the city of Amsterdam

making it relevant to the task, while the other is a general dataset. These datasets are used to train various seq2seq models, including T5-Small, BART-Base, mT5-Small, mBART-Large-50, T5-Base-Dutch, UL2-Small-Dutch and UL2-Small-Dutch-Simplification. Different approaches are employed, including training on individual datasets and sequential training on combined datasets. The performance of these models is evaluated using quantitative metrics such as the Flesch-Kincaid Grade Level (FKGL), BLEU score and SARI score. After which the sentences are briefly inspected qualitatively.

Following the evaluation, the best performing model is applied to the RVO letters to generate simplified versions. The letters are segmented into sentences for simplification and then re-assembled. Through this approach, text simplification is achieved by simplifying the individual sentences. The simplified letters are analyzed using FKGL scores and qualitatively checked to see if the original meaning is preserved.

2 Literature Review

2.1 Related Work

Text simplification (TS) is a vital area in natural language processing (NLP) aimed at making text more accessible by reducing its linguistic complexity while preserving the original meaning. This section reviews key studies in TS and related approaches using sequence-to-sequence models.

Bulté et al. (2018) examined the readability of Dutch texts and the potential for automated text simplification. Their study emphasized the importance of developing effective automated lexical tools to simplify such texts.

The project by Harmsen and Raaij (2023) focused on the automatic simplification of Dutch texts, specifically government communications. They evaluated various machine learning models, including rule-based and data-driven approaches, to determine their effectiveness in improving text comprehensibility. Their findings indicated that a models such as Naive and RobBERT are not suitable for the task, and while ChatGPT performed well, there is a transparency risk associated with ChatGPT, as it remains a black box model for humans (Wu et al., 2023).

Sheang and Saggion (2021) explored the use of a pre-trained seq2seq model T5 for text simplification, introducing a controllable mechanism to regulate system outputs. Their experiments show that T5, trained on several tasks with large a amount of data, can achieve remarkable results in text simplification.

The research by Alshantqiti et al. (2023) extended the application of seq2seq models to Arabic text simplification. Their work involved the use of a transformer-based seq2seq model with Fast Fourier Transform layers. Their results show that their model outperforms all state-of-the-art Arabic language text-to-text generation models. Which indicates that seq2seq transformer models can be extended to different languages to perform text simplification.

Seidl and Vandeghinste (2024) presented a supervised sentence simplification approach for Dutch using a pre-trained large language model (T5). Their investigation showed that although the approach did not outperform its own baseline, it did demonstrate the potential for sentence simplification tasks.

2.2 Theoretical Framework

2.2.1 Sequence-to-Sequence Model Architecture

Sequence-to-sequence models are a class of models designed to transform one sequence into another. They are particularly effective for tasks where the input and output are sequences of varying lengths, such as machine translation and text summarization. Introduced by Sutskever et al. (2014), these models consist of two main components: the encoder and the decoder.

The encoder processes the input sequence and transforms it into a fixed length context vector, effectively summarizing the information. This is typically achieved using Recurrent Neural Networks (RNNs) such as Long Short-Term Memory (LSTM) networks. The decoder then takes this context vector and generates the output sequence one element at a time. Similar to the encoder, the decoder is often an LSTM, predicting each token based on the context vector and the previously generated tokens. However, a challenge with seq2seq models is this reliance on a single context vector, which can limit performance on longer sequences. To address this, the attention mechanism was introduced. This allows the decoder to focus on different parts of the input sequence at each step, improving the handling of long and complex sequences.

2.2.2 Transformers

Transformers, introduced by Vaswani et al. (2017), are an advanced type of sequence-to-sequence model architecture. Instead of processing sequences one step at a time like traditional models, Transformers look at the whole sequence simultaneously using a mechanism called self-attention. This makes them faster and better at understanding context. Transformers have the same two main parts: the encoder, which reads and understands the input sequence, and the decoder, which generates the output sequence. Both parts use self-attention to focus on different parts of the sequence, capturing important relationships and meanings. Positional encoding is added to help the model understand the order of the tokens in the sequence.

Models like T5 (Text-to-Text Transfer Transformer) and BART (Bidirectional and Auto-Regressive Transformers) are based on the Transformer architecture and have achieved success in various natural language processing (NLP) tasks. T5 treats all NLP tasks as text-to-text problems, making it useful for tasks such as text classification, text generation, question answering and summarization. BART, with its bidirectional encoding and auto-regressive decoding, excels in tasks like text generation, summarization and translation. Both models leverage extensive pre-training on large datasets, enabling them to handle a wide range of NLP tasks (Weng, 2024).

3 RVO Dataset

3.1 Data Description

The dataset provided by the RVO contains PDF files, which were first converted to text format to process and analyze them effectively. The dataset consists of 131 Dutch letter templates, totaling 76,725 words, with an average of 586 words per letter. Basic statistics about the word counts of these letters can be seen in Table 1. These letters were sent by the RVO to entrepreneurs on a range of topics, such as the availability of funding programs and the eligibility of entrepreneurs to submit applications.

Statistic	Value
Count	131
Mean	586
Standard Deviation	553
Minimum	90
25th Percentile	280
50th Percentile (Median)	405
75th Percentile	696
Maximum	3439
Total Number of Words	76,725
Average Number of Words per Letter	586

Table 1: Basic Statistics (Pre-cleaning)

The distribution of letter lengths, as illustrated in Figure 1, indicates that most letters contain fewer than 1000 words, with some outliers exceeding 1500 words. The presence of a long tail in the distribution suggests the necessity of setting a threshold to effectively manage the outliers.

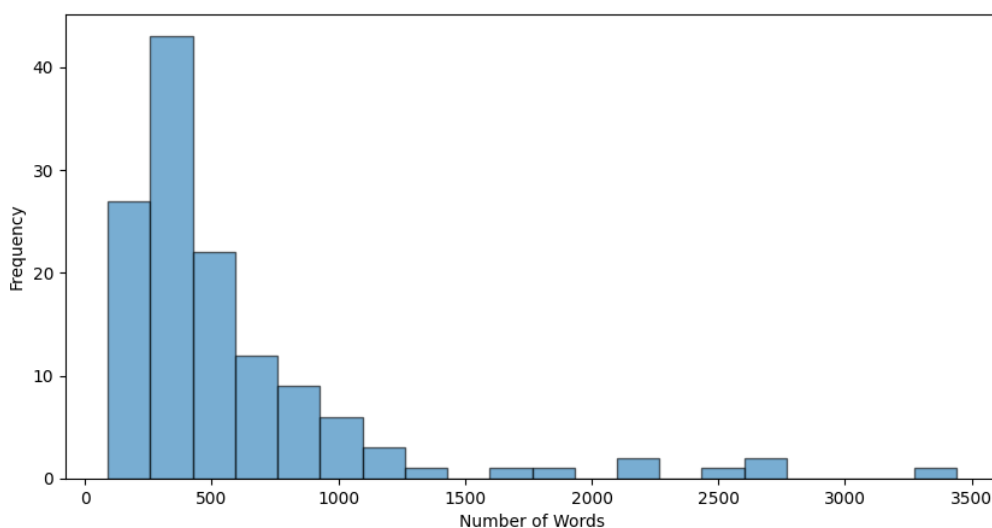


Figure 1: Histogram of Letter Lengths (Pre-cleaning)

3.2 Data Cleaning

First a threshold of 1500 words per letter was set. This decision was based on the analysis of the distribution of letter lengths, where it was observed that the majority of the letters are under 1000 words, with a few outliers exceeding 1500 words. This threshold ensures that the dataset is representative of typical letters while keeping the dataset within a reasonable size, making it more manageable for further processing and analysis.

As a result, eight letters exceed this threshold and were excluded from the dataset, leaving a total of 123 letters for analysis. The excluded letters and their word counts are listed in Table 2.

File Name	Word Count
61. UPEU Beslissing aanvraag Jonge Landbouwers Friesland	1636
54. UPEU Beslissing aanvraag Gezonde kalverket	1888
46. CITES Officiële waarschuwing Omgevingswet	2208
45. CITES Vooraankondiging bestuursrechtelijke handhaving	2231
11. Herbeschikking SVNL 2016	2524
47. CITES Bestuursdwangbesluit Omgevingswet	2612
43. CITES LOD OW 1	2726
44. CITES Uitleg omgevingswet	3439

Table 2: Excluded Letters

The data cleaning process was further conducted manually due to issues introduced during the conversion of PDF files to text format. The focus of the manual data cleaning process is on the preservation of the core content within the letters. The following steps were taken to prepare the dataset for text simplification:

- **Correction of Conversion Errors:** The conversion from PDF to text caused some elements of the letters to be incorrect. These errors were manually corrected to ensure the text maintains its intended structure and meaning.
- **Removal of Headers and Footers:** Data that is not relevant including return addresses, dates and page numbers were removed to focus on the core content of the letters.
- **Removal of Non-Content Text:** Irrelevant data including sender and recipient addresses, website URLs and contact information were removed. If such information is important for context, it was replaced with placeholders enclosed in brackets (e.g. [URL]).
- **Removal of Private Information:** Private information like passwords, codes and IBANs were removed due to privacy concerns. Similarly, email addresses, phone numbers and other sensitive data were either removed or replaced with placeholders if necessary for context.
- **Redundancy Removal:** Duplicate content within the letters was removed to prevent redundancy.
- **Exclusion of Attachments (Bijlages):** Attachments and appendices included in the letters were removed. These sections often contained information that were not central to the main content being analyzed for text simplification. Including these would add unnecessary complexity and could detract from the main focus of the analysis.

Furthermore, the letter "48. Wijziging standaard WIS" was removed from the dataset, as it only contained attachments.

After manually cleaning the dataset, 122 letters remained. Some basic words statistics from the letters can be seen in Table 3. The total number of words decreased from 76,725 to 43,191, while the average words per letter decreased from 586 to 354. No letters remained containing more than a 1000 words.

Statistic	Value
Count	122
Mean	354
Standard Deviation	202
Minimum	76
25th Percentile	217
50th Percentile (Median)	298
75th Percentile	442
Maximum	966
Total Number of Words	43,191
Average Number of Words per Letter	354

Table 3: Basic Statistics (Post-cleaning)

The histogram in Figure 2 shows the distribution of letter lengths after the data cleaning process, indicating a more uniform distribution without extreme outliers. The cleaned dataset supports more accurate and efficient text simplification, which should lead to better model performance.

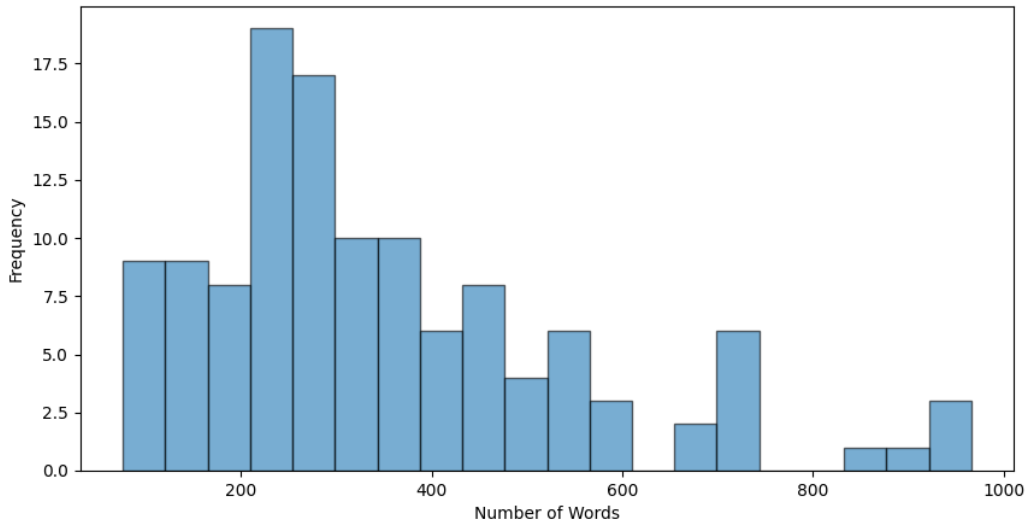


Figure 2: Histogram of Letter Lengths (Post-cleaning)

4 Methodology

4.1 Training and Evaluation Datasets

4.1.1 Municipal Dataset

The Municipal dataset is sourced from the Amsterdam AI Team and includes a variety of sentences relevant to governmental communications. The corpus consists of 1311 automatically aligned complex-simple sentence pairs. The original documents used to create the dataset were

provided by the Communications Department of the City of Amsterdam and include reports, citizen letters, newsletters and other types of documents. These documents were reviewed by an expert and they contain edits related to simplification, spelling corrections and other improvements. The relevance of the dataset lies in the direct alignment of making Dutch governmental letters more accessible.

The dataset was split into training (80%), validation (10%) and test (10%) sets resulting in the distribution seen in Table 4.

Data Set	Sentences
Training Set	1048
Validation Set	131
Test Set	131

Table 4: Municipal Dataset Sentence Split

4.1.2 UWV Dataset

The UWV dataset, also known as "veringewikkelderingen," is a general dataset designed to provide variations of Wikipedia paragraphs in different styles. The dataset consists of 1490 rows, where each row contains an original Wikipedia text and its simplified version. Unlike the Municipal dataset, the rows in the UWV dataset often contain more than one sentence, usually around two to three sentences. The dataset was generated using GPT-4 with a specific system prompt designed to create different versions of the input text using various styles, which include jargon, technical language and others. The original and simplified texts are the primary columns of interest for this study.

The dataset was split into training (80%), validation (10%) and test (10%) sets resulting in the distribution seen in Table 5.

Data Set	Sentences
Training Set	1192
Validation Set	149
Test Set	149

Table 5: UWV Dataset Sentence Split

4.2 Model Selection and Training

4.2.1 Models Selected

For this study, several seq2seq transformer models were chosen to evaluate their effectiveness in simplification, the specific models selected are expanded upon in the following paragraph.

T5-Small

The T5 model was introduced by Raffel et al. (2020), it reframes all NLP tasks into a unified text-to-text format where both input and output are text strings. The model achieves state-of-the-art results on multiple benchmarks including summarization, question answering, text classification and more. The T5-Small version is a more computationally efficient variant, which makes it suitable for environments with limited resources.

BART-Base

BART was developed by Lewis et al. (2019) and is known for its effectiveness in text generation tasks, such as summarization and translation. BART combines bidirectional and autoregressive transformers and its effectiveness in text generation makes it suitable for text simplification.

mT5-Small

mT5 is a multilingual variant of the T5 model and is designed to handle multiple languages effectively, introduced by Xue et al. (2020). mT5-Small is a scaled-down version that balances performance and computational efficiency, its multilingual training makes it a useful model for handling Dutch text simplification.

mBART-Large-50

mBART (Multilingual BART) is a multilingual extension of BART, pre-trained on multiple languages (Liu et al., 2020). Tang et al. (2020) developed mBART-Large-50 to handle text generation tasks across 50 different languages, leveraging its pre-training to perform well on Dutch text simplification.

UL2-Small-Dutch (UL2-SD)

UL2 (Unifying Language Learning Paradigms) introduced by Tay et al. (2022), is designed to address a variety of NLP tasks by unifying multiple language learning paradigms into a single framework. The UL2-SD version is fine-tuned specifically for Dutch and pre-trained on a large corpus of Dutch data.

UL2-Small-Dutch-Simplification (UL2-SDS)

UL2-SDS is a variant of the UL2-Dutch model that has been further fine-tuned specifically for the task of text simplification in Dutch. The model is intended for simplification of Dutch language on a sentence level.

4.2.2 Training Approaches

To evaluate the effectiveness of the selected seq2seq models the following strategies were employed.

Municipal Only: Models were trained exclusively on the Municipal dataset, which contained complex-simple sentence pairs from documents provided by the Communications Department of the City of Amsterdam.

UWV Only: Models were trained exclusively on the UWV dataset, which consisted of original-simplified sentence pairs derived from Wikipedia paragraphs.

Municipal then UWV: Models were first trained on the Municipal dataset and then further fine-tuned on the UWV dataset.

UWV then Municipal: Models were first trained on the UWV dataset and then further fine-tuned on the Municipal dataset.

The performance of these models was evaluated using FKGL, BLEU and SARI metrics. The best-performing model was used for simplifying the RVO letters.

4.2.3 Hyperparameter Tuning

Given the scope and computational limitations of this research, hyperparameter tuning was primarily based on settings derived from previous research, specifically the work by Seidl and Vandeghinste (2024) on Dutch sentence simplification. The settings used in their work were

adopted as-is for training with one single dataset and the initial training in the sequential training, using the AdamW optimizer, and can be seen in Table 6.

However, during the sequential training phases, the model’s performance did not improve as expected when fine-tuning with the second dataset. This observation suggested the need for further fine-tuning of hyperparameters. The learning rate was chosen as the parameter to alter and adjustments were explored within the specified range (10^{-3} to 10^{-5}).

A new learning rate of 5×10^{-5} was chosen, for fine-tuning with the second dataset, which is a moderate decrease from the initial rate. Preliminary evaluations with the new learning rate indicated improved model performance during sequential training, which justified its selection.

Parameter	Initial Training	Sequential Training
Seed	42	42
Gradient accumulation	4	4
Learning rate	0.0001	5×10^{-5}
Betas (β_1, β_2)	0.9, 0.999	0.9, 0.999
Batch size	6	6
Weight decay	0.1	0.1
Epochs	4	4
Warmup steps	5	5

Table 6: Hyperparameter Settings

4.3 Sentence Simplification

A sentence simplification procedure was developed to process the input text in batches, tokenize it and generate simplified output using the models. This involved processing the text in batches of size 8, to balance memory constraints and processing efficiency. Tokenizing the input text with padding and truncation to a maximum length of 512 tokens and generating simplified text using beam search with 4 beams and early stopping in order to generate higher quality outputs. This procedure was applied consistently across all models and training methods in order to maintain uniformity. All models and training methods were evaluated on the Municipal test set using the sentence simplification procedure. The models generated simplified sentences from the complex input sentences, which were then evaluated using the FKGL, BLEU and SARI metrics.

4.4 Evaluation Metrics

The performance of the seq2seq models and their training methods was evaluated using three primary metrics: FKGL (Flesch-Kincaid Grade Level), BLEU and SARI. Each of these metrics provides a different perspective on the quality and readability of the simplified texts. It is important to note that these metrics have some limitations, so a small qualitative check of the sentences was also conducted.

4.4.1 FKGL (Flesch-Kincaid Grade Level)

The FKGL metric is a measure that is often used for research regarding simplification (Seidl & Vandeghinste, 2024). It evaluates readability based on sentence length and word complexity and thereby provides a measure of the reading level required to understand the text (Flesch, 1948; Kincaid et al., 1975). The FKGL is calculated using the formula:

$$\text{FKGL} = 0.39 \times \left(\frac{\text{total words}}{\text{total sentences}} \right) + 11.8 \times \left(\frac{\text{total syllables}}{\text{total words}} \right) - 15.59$$

A lower FKGL score indicates that the text or sentence has higher readability. However, the metric is limited in the way it only looks at word and sentence length, and ignores grammaticality and semantic meaning (Wubben et al., 2012).

4.4.2 BLEU Score

Proposed by Papineni et al. (2002), the BLEU (Bilingual Evaluation Understudy) score is a metric that compares n-grams of the candidate text with reference texts. It is commonly used in machine translation and text generation tasks to measure how close a generated text is to human-written text. In this study, the BLEU score is computed by comparing the n-grams of the simplified sentences with reference simplified sentences. A higher BLEU score indicates a closer match to the reference text, reflecting better preservation of meaning and grammaticality.

BLEU score is limited because it does not always align well with simplicity when sentence splitting is performed (Alva-Manchego et al., 2020). However, BLEU does strongly correlate with human judgments of grammaticality and semantic meaning, making it a useful metric for evaluating the preservation of content in simplified text (Seidl & Vandeghinste, 2024).

4.4.3 SARI Score

SARI (System output Against References and against the Input sentence) is a metric introduced by Xu et al. (2016) to evaluate the quality of text simplification. Unlike BLEU, which focuses on n-gram precision, SARI specifically measures the goodness of words that are added, deleted, and kept by the system in comparison to the input sentence and reference simplifications. The metric is especially useful for text simplification because it was designed to capture the specific operations involved in simplifying text. However, the effectiveness of the metric decreases when only a single reference sentence is available. A higher SARI score indicates better text simplification performance by the system, as it means the system is adding, deleting and keeping the correct words in comparison to the input sentence and reference simplifications.

4.5 Rewriting RVO Letters

4.5.1 Segmentation and Simplification of Letters

The best-performing model, identified through comparing the evaluation metrics as well as a small qualitative check for any inconsistencies or notabilities, was applied to the RVO letters to generate simplified versions. Each letter was first segmented into individual sentences using the SpaCy library. The same sentence simplification procedure used for the test sets was then applied to these segments.

4.5.2 Evaluation of Simplified Letters

The simplified letters were then reassembled from the simplified sentences. The readability of the simplified letters was then evaluated using the FKGL metric, incorporating BLEU and SARI was not possible, as reference letters were not provided. Additionally, a qualitative inspection was conducted to verify that the simplified text maintained the information and context of the original letters.

4.6 Implementation

All models were trained and implemented using the Hugging Face Transformers library (Wolf et al., 2020). The Textstat library was used to calculate the Flesch-Kincaid Grade Level (FKGL) metric, SacreBLEU was used to compute the BLEU score, the Evaluate library was used for calculating the SARI score and spaCy was used to segment the letters into individual sentences.

The computational resources for this study were provided by Google Colab Pro. The training and evaluation were conducted using an NVIDIA T4 GPU with a high RAM setting. The environment was configured with Python 3.10.12 and additional dependencies included pandas, NumPy, scikit-learn, Torch, Matplotlib and Seaborn.

5 Results

5.1 Performance of Models on Test Set

The performance metrics for each model and training combination are shown in Table 7. The models evaluated include T5-Small, BART-Base, mT5-Small, mBART-Large-50, UL2-Small-Dutch (UL2-SD), and UL2-Small-Dutch-Simplification (UL2-SDS). The training approaches include training exclusively on the Municipal dataset, exclusively on the UWV dataset, sequential training (Municipal then UWV), and sequential training (UWV then Municipal).

Model	Training Approach	FKGL ↓	BLEU ↑	SARI ↑
T5-Small	Municipal Only	9.00	52.42	53.61
T5-Small	UWV Only	8.95	50.08	52.02
T5-Small	Municipal then UWV	9.09	50.74	50.76
T5-Small	UWV then Municipal	8.93	52.72	54.86
BART-Base	Municipal Only	8.48	53.78	53.85
BART-Base	UWV Only	7.03	32.96	37.78
BART-Base	Municipal then UWV	7.00	38.41	42.03
BART-Base	UWV then Municipal	8.48	52.82	53.05
mT5-Small	Municipal Only	8.77	1.39	17.66
mT5-Small	UWV Only	-	-	-
mT5-Small	Municipal then UWV	-	-	-
mT5-Small	UWV then Municipal	-	-	-
mBART-Large	Municipal Only	8.13	55.27	58.05
mBART-Large	UWV Only	-	-	-
mBART-Large	Municipal then UWV	-	-	-
mBART-Large	UWV then Municipal	-	-	-
UL2-SD	Municipal Only	9.46	22.81	39.90
UL2-SD	UWV Only	7.93	21.28	36.43
UL2-SD	Municipal then UWV	8.28	18.59	36.16
UL2-SD	UWV then Municipal	8.47	42.16	44.02
UL2-SDS	Municipal Only	8.59	43.02	45.02
UL2-SDS	UWV Only	7.59	18.59	39.50
UL2-SDS	Municipal then UWV	8.18	26.36	40.11
UL2-SDS	UWV then Municipal	8.66	42.50	45.79

Table 7: Performance Metrics for Each Model and Training Combination

↓ indicates that lower is better, ↑ indicates that higher is better.

Note: The mT5-Small and mBART-Large-50 models encountered CUDA memory issues during training on the UWV dataset, which is why these models were only trained on the Municipal dataset.

5.2 Comparative Analysis

5.2.1 Quantitative Analysis

The original FKGL score of the sentences is 9.22 and serves as a baseline. The models that are chosen for further qualitative evaluation, according to their performance metrics in the best training strategy for that model, are summarized in Table 8. The notable exception being the mT5-Small model, which scored extremely low in BLEU and SARI and is therefore not evaluated further. The selection focuses on the lowest FKGL scores, provided that the SARI and BLEU scores are high or competitive with the other training strategies for that model.

Model	Training Strategy	FKGL (Original) ↓	FKGL ↓	BLEU ↑	SARI ↑
T5-Small	UWV then Municipal	9.22	8.93	52.72	54.86
BART-Base	Municipal Only	9.22	8.48	53.78	53.85
mBART-Large	Municipal Only	9.22	8.13	55.27	58.05
UL2-SD	UWV then Municipal	9.22	8.47	42.16	44.02
UL2-SDS	Municipal Only	9.22	8.59	43.02	45.02

Table 8: Summary of Best Scores

- **T5-Small (UWV then Municipal):** Chosen for the highest scores across all metrics, best FKGL (8.93), BLEU (52.72) and SARI (54.86).
- **BART-Base (Municipal Only):** Selected because the FKGL (8.48) score is lower than baseline, high BLEU (53.78) and high SARI (53.85) scores.
- **mBART-Large (Municipal Only):** Only training strategy for this model, included as it scores a competitive FKGL (8.13), and the highest BLEU (55.27) and SARI (58.05) scores for all model and training combinations.
- **UL2-SD (UWV then Municipal):** Included as it has FKGL (8.47) score below baseline, with very competitive BLEU (42.16) and SARI (44.02) scores.
- **UL2-SDS (Municipal Only):** Chosen for better FKGL (8.59) over sequential training strategy (8.66), with competitive BLEU (43.02) and SARI (45.02) scores.

5.2.2 Qualitative Analysis

For the qualitative evaluation, simplified sentences were compared to the original and reference sentences from the Municipal dataset. Although all sentences were rewritten to calculate the metrics, this section focuses on three representative examples. Based on the characteristics and notable aspects of how these sentences are simplified, a model is chosen to simplify the RVO letters. It should be noted that while only three examples are shown here, more sentences were checked but are not included in the text for brevity.

This qualitative reveals strengths or weaknesses not evident in the quantitative metrics alone. Therefore, it is possible that the best model identified through this qualitative analysis might not align perfectly with the quantitative results.

Table 9 shows the first example of sentence simplification using the models. In this example, both T5-Small and UL2-SD repeated the original sentence verbatim, which indicates a failure to simplify. BART-Base produced an incomplete and incorrect sentence, rendering it unusable.

mBART-Large provided a slight simplification by adjusting the sentence structure, but produced an incorrect sentence. UL2-SDS also retained much of the original structure, showing a minimal increase in readability.

Sentence Type	Sentence
Original	Als de zaak een zaaknummer heeft moet er nog een documentnummer worden aangevraagd.
Reference	De zaak heeft al een zaaknummer. Je moet nu nog een documentnummer aanvragen.
T5-Small	Als de zaak een zaaknummer heeft moet er nog een documentnummer worden aangevraagd.
BART-Base	We moet er nog een documentnummer aangevraagd.
mBART-Large	Als de zaak een zaaknummer heeft, moet er nog een documentnummer aanvragen.
UL2-SD	Als de zaak een zaaknummer heeft moet er nog een documentnummer worden aangevraagd.
UL2-SDS	Als de zaak een zaaknummer heeft, moet er nog een documentnummer worden aangevraagd.

Table 9: Simplifications of Example Sentence 1

In the second example, as seen in table 10, T5-Small again failed to simplify, repeating the original sentence. BART-Base attempted a simplification but introduced grammatical errors. mBART-Large made slight changes but resulted in awkward phrasing. Both UL2-SD and UL2-SDS provided effective simplifications by rephrasing the sentence while retaining its meaning, although the original structure was largely maintained.

Sentence Type	Sentence
Original	Hierbij wordt een beperking beschouwd als een afwijking die door bijvoorbeeld dokters of hulpverleners weggenomen moet worden.
Reference	Dan zien we een beperking als een afwijking die bijvoorbeeld dokters of hulpverleners weg moeten nemen.
T5-Small	Hierbij wordt een beperking beschouwd als een afwijking die door bijvoorbeeld dokters of hulpverleners weggenomen moet worden.
BART-Base	Hierbij wordt een beperking beschouwd als een afwijking die door bijvoorbeeld dokters of hulpverleners weggenomen moeten.
mBART-Large	Hierbij wordt een beperking gezien als een afwijking die bijvoorbeeld dokters of hulpverleners wegnemen moeten.
UL2-SD	Dit is een beperking die door bijvoorbeeld dokters of hulpverleners weggenomen moet worden.
UL2-SDS	Dit is een beperking die door bijvoorbeeld dokters of hulpverleners weggenomen moet worden.

Table 10: Simplifications of Example Sentence 2

In the third example, T5-Small repeated the original sentence again, showing no simplification. BART-Base simplified the sentence but omitted "geen" changing the meaning. mBART-Large closely matched the reference sentence, maintaining readability and meaning. UL2-SD introduced errors, while UL2-SDS effectively simplified the sentence but altered the meaning by adding "nodig."

Sentence Type	Sentence
Original	U heeft mogelijk geen (geldig) identiteitsbewijs
Reference	U hebt misschien geen (geldig) identiteitsbewijs
T5-Small	U heeft mogelijk geen (geldig) identiteitsbewijs
BART-Base	U hebt mogelijk (geldig) identiteitsbewijs
mBART-Large	U hebt misschien geen (geldig) identiteitsbewijs
UL2-SD	U heeft mogelijk geen (geldig identiteitsbewijs)bewijs.
UL2-SDS	U hebt waarschijnlijk geen (geldig) identiteitsbewijs nodig.

Table 11: Simplifications of Example Sentence 3

T5-Small frequently repeated the original sentences verbatim, indicating an issue with over-reliance on the input and a failure to simplify effectively. This tendency diminishes its usefulness for text simplification tasks, as seen in the examples where it failed to produce meaningful simplifications. While BART-Base attempted simplification, it often introduced grammatical errors or produced incomplete sentences. These issues reduce the readability and accuracy of the output, which makes BART-Base a less reliable option for generating simplified text that retains the original meaning. mBART-Large, provided the closest matches to the reference sentences, maintaining readability and meaning effectively. However, it sometimes introduced awkward phrasing or failed to simplify as thoroughly as desired. Despite these issues, mBART-Large generally performed well in producing simplified text. UL2-SD showed potential but introduced grammatical errors, which compromise the readability and accuracy of the output. Although it simplified some sentences effectively, the inconsistency in performance makes it less reliable for widespread use. UL2-SDS effectively simplified sentences while often preserving the original meaning and improving readability. However, at times it did alter the meaning slightly as seen in one of the examples.

5.3 Model Selection for RVO Simplification

Based on both quantitative and qualitative analyses, UL2-SDS (Municipal Only) is chosen to perform the text simplification for the RVO letters. Although the model had slightly lower quantitative scores compared to some others, it showed the least problems in the qualitative analysis. Additionally, UL2-SDS was specifically trained and created for Dutch simplification, which makes it the most promising option. For models that appeared fine in the qualitative analysis and had similarly high metric scores, any of them could have been chosen. However, a decision had to be made, and given UL2-SDS’s training for simplification tasks, it is selected as the best fit.

5.4 Evaluation of Simplified RVO Letters

5.4.1 Quantitative: FKGL Scores

The evaluation of the simplified RVO letters includes both quantitative and qualitative analyses to assess the effectiveness of the simplification process using the UL2-SDS model trained with the municipal dataset only. The distribution of the FKGL scores for the original letters can be seen in Figure 3.

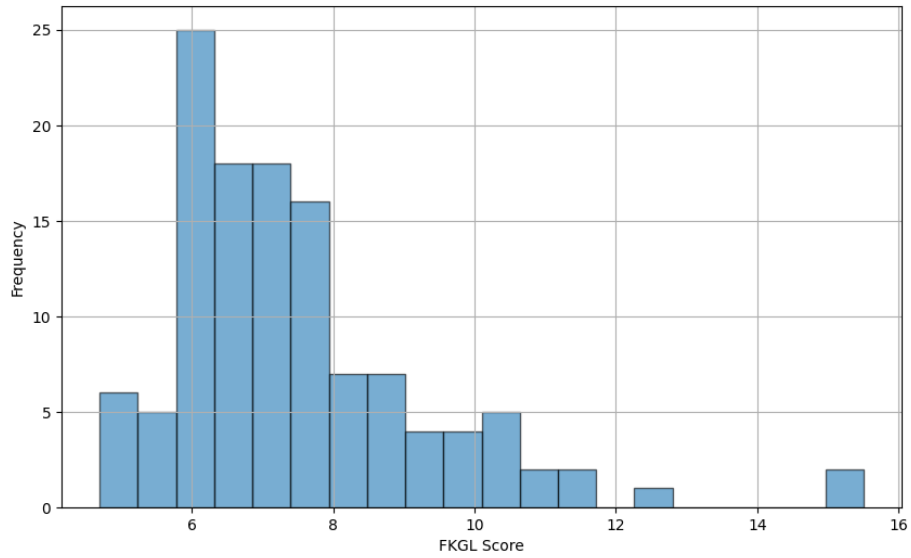


Figure 3: Histogram of FKGL Scores for Original Letters

The average FKGL score for the original RVO letters is 7.50, while the average FKGL score for the simplified letters is 7.11. This reduction in FKGL score indicates an improvement in readability, which suggests that the simplification process has made the text easier to understand. This is further supported by the median FKGL score reducing from 7.1 for the original RVO letters, to 6.8 for the simplified letters. The distribution of the FKGL scores for the rewritten letters are shown in Figure 4.

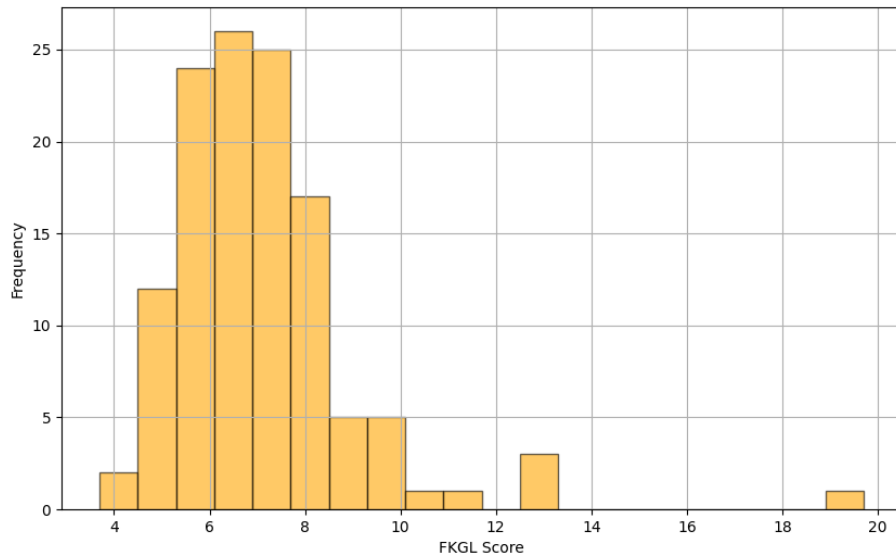


Figure 4: Histogram of FKGL Scores for Simplified Letters

The density plot of the FKGL scores, shown in Figure 5, provides another perspective on the distribution of readability scores. The density plot reveals a clear peak at lower FKGL scores

for the simplified letters compared to the original letters, which indicates that the simplified letters are consistently more readable, as the density of lower FKGL scores is higher for the simplified texts.

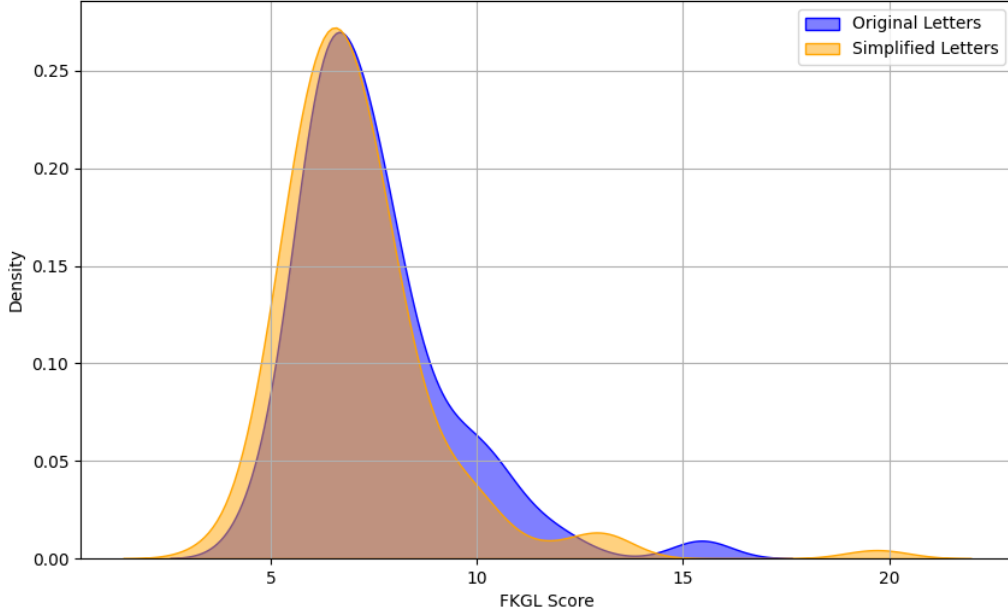


Figure 5: Density Plot of FKGL Scores for Original and Simplified Letters

5.4.2 Qualitative: Inspecting Rewritten Letters

To gain deeper insights into the effectiveness of the model, a few letters were manually reviewed. The selection of these letters was based on a range of FKGL scores, in order to create a representative sample that included both outliers and letters with scores close to the average. The full original and simplified letters are included in the appendix to provide detailed context.

Highest FKGL Scores

The simplified version of the letter "09. Terugbetaling heffing DGF" indicates a potential issue with the rewriting process. The FKGL score increased significantly from 8.2 to 19.7, indicating a serious decrease in readability. The repetition of the word "Terugbetaling" throughout the text is clearly an error, making the letter nonsensical and difficult to read. The simplified letter also contains grammatical errors and incomplete sentences, further decreasing the readability of the text. The most problematic point in this letter, is the error propagation in the model by repeating one word excessively. According to Zhang et al. (2023), seq2seq models are prone to generating redundant repetitions. The problem arises because the models tend to assign higher probabilities to repetitive and generic sequences over more diverse and natural text. This repetition can also occur due to the model's over-attention to certain tokens, as it fails to properly track which source tokens have already been covered in past attentions, leading to repeated focus on the same tokens. In this case, the excessive repetition of "Terugbetaling" could result from such an over-attention issue. This issue, while significant in this instance, could be an outlier as this letter is the only one with such a drastically high FKGL score.

In order to see if this is truly the case, the rewritten letters with the ten highest FKGL scores were checked. Several letters also exhibited problematic or excessive repetition. For example,

letters such as "51. Ontvangstbevestiging aanvraag vaststelling subsidie Installaties voor lagere stikstofuitstoot garnalenvissers" (Rank 2), "14. Bewijs van inschrijving" (Rank 3), "08. Herinnering heffing DGF" (Rank 6), and "80. Overdracht SABE Advies over duurzame landbouw voor agrariërs" (Rank 9) showed similar issues with repeated phrases or tokens. However, not all of these letters had repetition issues to the same severity, especially in comparison to the letter "09. Terugbetaling heffing DGF" which had the most drastic increase in FKGL score and the most severe repetition errors. This suggests that while the model generally performs well, it can encounter significant failures with specific inputs.

Average FKGL Scores

Letters with FKGL scores closest to the mean were also analyzed, as these letters represent what the average rewritten looks like, some notabilities of these letters are discussed in this paragraph. A significant problem that arose again was the excessive repetition of phrases in letters like "37. Aanvullende gegevens landbouwer" and "38. Aanvullende gegevens landbouwer mineralenconcentraat" For instance, "Aanvullende gegevens landbouwer 2023" was repeated multiple times in both of these letters, this shows that the problem is also present in the average letter and not only in the letters with a high FKGL score.

The handling of placeholders indicated another area for improvement. In "97. UPnl Toewijzen aanvraag RUS2-3 EG-22 Geredigeerd" placeholders like [Bedrag] and [URL] were often replaced with generic tokens or left as placeholders in the simplified version, such as "TSBedrag" and "SCHURLURL" This impacted the coherence and usability of the letters, making them confusing and decreasing readability.

Despite these issues critical information and instructions were generally preserved. Essential details like submission deadlines and procedural steps were maintained across the simplified versions. However, the introduction of incomplete sentences and placeholders detracted from the overall quality. For example, in the simplified letter "97. UPnl Toewijzen aanvraag RUS2-3 EG-22 Geredigeerd" the original phrase "Ik heb uw aanvraag ontvangen voor EG-22-[Projectnummer]" was poorly rewritten to "Ik heb uw aanvraag ontvangen voor EG-22- 2022" failing to accurately replace the placeholder.

Biggest FKGL Score Improvements

In the case of "05. Factuur Binnenvisserij" the FKGL score improved from 15.4 to 9.1, which is an extreme increase in readability. While some sentences in the rewritten letters do seem easier to read, it is also clear that crucial details, such as the name of a relevant bank account, were omitted. In "30. Factuur Grondkamer Zuidwest" the FKGL score dropped from 15.5 to 11.5. The simplified version retained the essential details about the invoice and payment instructions but mishandled placeholders. For example, placeholders like [IBAN] and [Relatienummer] were either replaced with generic terms or completely left out, leading to instructions that lacked critical information, such as "de afdeling klantcontact, telefoon LDRelatienummer." The handling of placeholders was a recurring issue. In "65. UPnl Afwijzen aanvraag inhoudelijk SPUK" the FKGL score improved from 9.4 to 6.2. However the same issues arose, placeholders like [Projectnaam] were incorrectly simplified to nonsensical terms like "SCHuk iw 2022" which could confuse the reader. Moreover, certain terms were again repeated multiple times.

6 Discussion

6.1 Interpretation of the Results

The results of the quantitative part of the study demonstrate that the readability of Dutch governmental letters was improved using seq2seq models. The average FKGL score of the original

letters was 7.50, which decreased to 7.11 after simplification. This reduction in FKGL score suggests that the letters became easier to read. However, it is important to note that the FKGL metric primarily considers word and sentence length and does not account for grammaticality or semantic meaning.

In addition to the quantitative analysis, a qualitative evaluation was conducted to gain deeper insights into the simplification process. The qualitative analysis revealed several issues with the simplified letters. One such issue was the excessive repetition of words or phrases, which often rendered the text nonsensical. For example, in the letter "09. Terugbetaling heffing DGF," the word "Terugbetaling" was repeated very frequently, leading to an FKGL score of 19.7 in the simplified version, compared to 8.2 in the original. This repetition issue indicates a limitation of the seq2seq models for text simplification, as they sometimes fail to maintain coherence while rewriting the text. Another issue observed in the qualitative analysis was the improper handling of placeholders and private information. Simplified letters sometimes retained placeholders such as "[Bedrag]" or replaced them with nonsensical tokens like "TSBedrag." This issue impacted the readability and usefulness of the letters, as critical information necessary for understanding the context was lost.

6.2 Implications for Readability Improvement

The implications of these findings are twofold. Firstly, there is a clear potential for seq2seq models to aid in simplifying official documents, making them more readable for the general public. This could improve the effectiveness of communication from government agencies, which can help in making important information accessible to citizens.

Secondly, the identified qualitative issues underscore the need for further refinement of these models. Addressing these limitations is critical to making the simplified texts not only easier to read but also to preserve their original meaning and context. This shows the importance of combining automated solutions with additional quality control measures to achieve reliable text simplification.

6.3 Limitations of the Study

The conducted research contains several limitations which must be acknowledged. Firstly, the study primarily relied on the FKGL metric to evaluate readability improvements. Although FKGL provides a useful measure of sentence and word length, it does not consider grammaticality or the meaning of text. This limitation means that while the simplified texts may appear easier to read based on FKGL scores, they may not necessarily be more understandable or contextually accurate. Secondly, the qualitative analysis revealed significant issues with the simplified texts, including excessive repetition of words and improper handling of placeholders. These issues suggest that the seq2seq model sometimes fail to actually rewrite original text into correct and usable simplified text.

Another limitation is related to the datasets used for training and evaluation. The study used datasets derived from municipal communications and Wikipedia paragraphs, which may not fully represent the variability of all governmental letters. This could limit the generalizability of the findings to other types of official documents that have different structures. In addition, the UWV dataset was split into training, validation and test sets, but the test set was not utilized in the evaluation.

During the training process, the mT5-Small and mBART-Large-50 models encountered CUDA memory issues when trained on the UWV dataset. This led to an interruption in their training, and as a result, performance metrics for these models using the UWV dataset are unavailable. The memory issues may be attributed to the larger size and complexity of the UWV dataset,

which required more GPU memory than available. Moreover, the study’s approach to model training and hyperparameter tuning was based on settings from previous research. While this method provides a standardized starting point, it may not have been optimal for the specific task of training models to simplify RVO letters.

Lastly, the evaluation of the simplified letters lacked the incorporation of additional readability and quality metrics such as BLEU and SARI due to the absence of reference texts. This omission might have provided a better understanding of the models’ performance in maintaining the original content’s meaning.

6.4 Suggestions for Future Research

Firstly, addressing the issues of excessive repetition and improper handling of placeholders is of importance. Developing models that can better manage these aspects will make sure that the simplified texts are both coherent and retain their original meaning. Techniques such as improved attention mechanisms or post-processing steps could be explored to mitigate these problems.

Secondly, expanding the diversity and representativeness of training datasets could prove useful. Incorporating a wider range of governmental documents, including those with different structures and complexities, can help improve the generalizability of the models. In addition, creating reference letters for the original letters could allow for better training and the incorporation of additional evaluation metrics, such as BLEU and SARI, alongside FKGL, which will allow for better a understanding for model performance.

Thirdly, addressing the technical limitations encountered during model training, such as CUDA memory issues, is necessary. Exploring more efficient model architectures or using resources with higher computational capabilities can help overcome these challenges and enable the training of larger and more complex datasets.

Finally, exploring advanced seq2seq models that incorporate control tokens can address the limitations identified in this study, such as excessive repetition and improper handling of placeholders. Control tokens allow for fine-tuning various aspects of text simplification, such as sentence length, word length, paraphrasing, and lexical and syntactic complexity, which can improve the simplification performance.

7 Conclusion

This study aimed to answer the research question: To what extent can sequence-to-sequence transformer models improve the readability of Dutch official letters provided by the Rijksdienst voor Ondernemend Nederland?

The findings indicate that while seq2seq models can modestly enhance the readability of these letters, the improvement is relatively minor. The average FKGL score of the original letters decreased slightly from 7.50 to 7.11 after simplification using the best-performing model, UL2-SDS, fine-tuned on only the municipal dataset. Despite this reduction, the qualitative analysis revealed issues such as excessive repetition and mishandling of placeholders, which can affect the coherence and overall quality of the simplified text.

By applying seq2seq models, it is observed that these models hold some potential for making governmental communications more accessible. However, the improvements in readability were not as substantial as hoped, indicating the need for further refinement of these models to better handle the complexities of text and sentence simplification.

In conclusion, this research highlights the potential and limitations of using seq2seq models for text simplification of Dutch governmental letters. While there are modest gains in readability, significant challenges remain, particularly in maintaining the meaning of text and handling specific content elements accurately. Future research should focus on enhancing model performance to achieve more substantial improvements in simplification. This includes addressing the identified issues and exploring additional quality control measures to ensure that simplified texts are both readable and coherent.

References

- Alkaldi, W. (2022). *Enhancing text readability using deep learning techniques* [Doctoral dissertation, Université d’Ottawa/University of Ottawa].
- Alshancqiti, A., Alkhodre, A., Namoun, A., Albouq, S., & Nabil, E. (2023). A transformer seq2seq model with fast fourier transform layers for rephrasing and simplifying complex arabic text. *International Journal of Advanced Computer Science and Applications*, 14(2).
- Al-Thanyyan, S. S., & Azmi, A. M. (2021). Automated text simplification: A survey. *ACM Computing Surveys (CSUR)*, 54(2), 1–36.
- Alva-Manchego, F., Scarton, C., & Specia, L. (2020). Data-driven sentence simplification: Survey and benchmark. *Computational Linguistics*, 46(1), 135–187.
- Bulté, B., Sevens, L., & Vandeghinste, V. (2018). Automating lexical simplification in dutch. *Computational Linguistics in the Netherlands Journal*, 8, 24–48.
- Flesch, R. (1948). A new readability yardstick. *Journal of applied psychology*, 32(3), 221.
- Harmsen, F., & Raaij, N. V. (2023, September). *Automatic simplification of dutch texts* (tech. rep.) (Project Report ELSA lab versie 2.0 - ENG). BSSC. <https://www.brightlands.com/sites/default/files/2023-09/BSSC%20-%20Project%20Report%20ELSA%20lab%20versie%202.0%20-%20ENG.pdf>
- Hobo, E., Pouw, C., & Beinborn, L. (2023). “geen makkie”: Interpretable classification and simplification of dutch text complexity. *Proceedings of the 18th Workshop on Innovative Use of NLP for Building Educational Applications (BEA 2023)*, 503–517.
- Kincaid, J. P., Fishburne Jr, R. P., Rogers, R. L., & Chissom, B. S. (1975). Derivation of new readability formulas (automated readability index, fog count and flesch reading ease formula) for navy enlisted personnel.
- Lewis, M., Liu, Y., Goyal, N., Ghazvininejad, M., Mohamed, A., Levy, O., Stoyanov, V., & Zettlemoyer, L. (2019). Bart: Denoising sequence-to-sequence pre-training for natural language generation, translation, and comprehension. *arXiv preprint arXiv:1910.13461*.
- Liu, Y., Gu, J., Goyal, N., Li, X., Edunov, S., Ghazvininejad, M., Lewis, M., & Zettlemoyer, L. (2020). Multilingual denoising pre-training for neural machine translation. *Transactions of the Association for Computational Linguistics*, 8, 726–742.
- Martin, L., Fan, A., De La Clergerie, É. V., Bordes, A., & Sagot, B. (2022). Muss: Multilingual unsupervised sentence simplification by mining paraphrases. *Proceedings of the Thirteenth Language Resources and Evaluation Conference*, 1651–1664.
- Papineni, K., Roukos, S., Ward, T., & Zhu, W.-J. (2002). Bleu: A method for automatic evaluation of machine translation. *Proceedings of the 40th annual meeting of the Association for Computational Linguistics*, 311–318.
- Raffel, C., Shazeer, N., Roberts, A., Lee, K., Narang, S., Matena, M., Zhou, Y., Li, W., & Liu, P. J. (2020). Exploring the limits of transfer learning with a unified text-to-text transformer. *Journal of machine learning research*, 21(140), 1–67.
- Rijksdienst voor Ondernemend Nederland. (2024). About us. <https://english.rvo.nl/topics/about-us>
- Seidl, T., & Vandeghinste, V. (2024). Controllable sentence simplification in dutch. *Computational Linguistics in the Netherlands Journal*, 13, 31–61.
- Shardlow, M. (2014). A survey of automated text simplification. *International Journal of Advanced Computer Science and Applications*, 4(1), 58–70.
- Sheang, K. C., & Saggion, H. (2021). Controllable sentence simplification with a unified text-to-text transfer transformer. *Proceedings of the 14th International Conference on Natural Language Generation (INLG); 2021 Sep 20-24; Aberdeen, Scotland, UK. Aberdeen: Association for Computational Linguistics; 2021*.
- Štajner, S. (2021). Automatic text simplification for social good: Progress and challenges. *Findings of the Association for Computational Linguistics: ACL-IJCNLP 2021*, 2637–2652.

- Sutskever, I., Vinyals, O., & Le, Q. V. (2014). Sequence to sequence learning with neural networks. *Advances in neural information processing systems*, 27.
- Tang, Y., Tran, C., Li, X., Chen, P.-J., Goyal, N., Chaudhary, V., Gu, J., & Fan, A. (2020). Multilingual translation with extensible multilingual pretraining and finetuning. *arXiv preprint arXiv:2008.00401*.
- Tay, Y., Dehghani, M., Tran, V. Q., Garcia, X., Wei, J., Wang, X., Chung, H. W., Shakeri, S., Bahri, D., Schuster, T., et al. (2022). U12: Unifying language learning paradigms. *arXiv preprint arXiv:2205.05131*.
- Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł., & Polosukhin, I. (2017). Attention is all you need. *Advances in neural information processing systems*, 30.
- Weng, B. (2024). Navigating the landscape of large language models: A comprehensive review and analysis of paradigms and fine-tuning strategies. *arXiv preprint arXiv:2404.09022*.
- Wolf, T., Debut, L., Sanh, V., Chaumond, J., Delangue, C., Moi, A., Cistac, P., Rault, T., Louf, R., Funtowicz, M., et al. (2020). Transformers: State-of-the-art natural language processing. *Proceedings of the 2020 conference on empirical methods in natural language processing: system demonstrations*, 38–45.
- Wu, T., He, S., Liu, J., Sun, S., Liu, K., Han, Q.-L., & Tang, Y. (2023). A brief overview of chatgpt: The history, status quo and potential future development. *IEEE/CAA Journal of Automatica Sinica*, 10(5), 1122–1136.
- Wubben, S., Van Den Bosch, A., & Krahmer, E. (2012). Sentence simplification by monolingual machine translation. *Proceedings of the 50th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, 1015–1024.
- Xu, W., Napoles, C., Pavlick, E., Chen, Q., & Callison-Burch, C. (2016). Optimizing statistical machine translation for text simplification. *Transactions of the Association for Computational Linguistics*, 4, 401–415.
- Xue, L., Constant, N., Roberts, A., Kale, M., Al-Rfou, R., Siddhant, A., Barua, A., & Raffel, C. (2020). Mt5: A massively multilingual pre-trained text-to-text transformer. *arXiv preprint arXiv:2010.11934*.
- Zhang, Y., Kamigaito, H., Aoki, T., Takamura, H., & Okumura, M. (2023). Generic mechanism for reducing repetitions in encoder-decoder models. *Journal of Natural Language Processing*, 30(2), 401–431.

Betreft: Ontvangstbevestiging aanvraag vaststelling subsidie Installaties voor lagere stikstofuitstoot garnalenvissers

Geachte heer/mevrouw,

Wij hebben uw aanvraag voor vaststelling van de subsidie Installaties voor lagere stikstofuitstoot garnalenvissers ontvangen op 20 november 2023 om 12:39:59 uur. Het referentienummer dat wij hieraan hebben gegeven is [Referentienummer]. Gebruik dit nummer als u met ons belt of digitale post wilt versturen.

Uw gegevens Naam: [Naam] KvK-nummer: [KvK-nummer] Relatienummer: [Relatienummer]

Meer informatie Heeft u vragen, kijk dan op [URL]. Of neem telefonisch contact met ons op: [Telefoonnummer] (lokaal tarief).

Met vriendelijke groet,

De Minister van Landbouw, Natuur en Voedselkwaliteit

Rewritten FKGL: 13.1

Rewritten Content:

We hebben uw aanvraag voor vaststelling van de subsidie voor lagere stikstofuitstoot garnalenvissers ontvangen op 20 november 2023 om 12:39:59 uur. Het referentienummer dat wij hieraan geven is ODReferentienummer. Dit nummer als u met ons belt of digitale post wilt versturen. KvK-nummer: LDRelatienummer: LDRelatienummer: LDRelatienummer: LDRelatienummer: LDRelatienummer: LDRelatienummer: LDRelatienummer: LDRelatienummer: LDRelatienummer Of neem telefonisch contact met ons op: ODTelefoonnummer LD (lokaal tarief). De Minister van Landbouw, Natuur en Voedselkwaliteit groet, De Minister van Landbouw, Natuur en Voedselkwaliteit.

Letter 3: 14. Bewijs van inschrijving (FKGL Score: 3rd highest)

Original FKGL: 11.0

Original Content:

Betreft: Bewijs van inschrijving

Geachte heer/mevrouw,

U krijgt dit Bewijs van inschrijving omdat u voor het eerst zaken met ons doet, of omdat u dit bewijs heeft aangevraagd. Hieronder ziet u de gegevens die op dit moment bij ons zijn geregistreerd.

Inloggen met eHerkenning voor ondernemers en organisaties Als ondernemer of organisatie logt u op [URL] in via eHerkenning. eHerkenning is de digitale toegangssleutel die u steeds vaker nodig zult hebben om bij overheidsinstellingen in te kunnen loggen. Om eHerkenning te kunnen gebruiken, heeft u een eHerkenningmiddel met betrouwbaarheidsniveau 2+ nodig. Zo'n middel kunt u aanschaffen bij één van de erkende aanbieders. Zie voor meer informatie op [URL].

Basisgegevens KVK-nummer: [KVK-nummer] Relatienummer: [Relatienummer] Naam: [Naam] Adres: [Adres] Oprichtingsdatum: 13 april [Jaar]

Gegevens voorkeursvestiging Vestigingsnummer: [Vestigingsnummer] Handelsnaam vestiging: [Handelsnaam vestiging] Correspondentieadres: [Adres] Telefoonnummer: [Telefoonnummer] E-mail: [Emailadres]

Gegevens wijzigen U kunt uw gegevens wijzigen of aanvullen via ons klantportaal ([URL]). Kies, nadat u heeft ingelogd, voor Mijn dossier en vervolgens voor Mijn gegevens wijzigen.

Denk hierbij aan het doorgeven van uw rekeningnummer als u subsidies wilt aanvragen of het doorgeven van uw e-mailadres zodat wij u digitaal serviceberichten kunnen sturen.

Meer informatie Heeft u vragen, kijk dan op [URL]. Of neem telefonisch contact met ons op: [Telefoonnummer] (lokaal tarief).

Met vriendelijke groet,

De Minister van Landbouw, Natuur en Voedselkwaliteit

Rewritten FKGL: 13.1

Rewritten Content:

Betreffen: Betreffen: Betreffen: Betreffen: Betreffen: Betreffen: Betreffen: Betreffen: Betreffen: Betreffen: Betreffen: Betreffen: Betreffen: Bewijs van inschrijving Dit overzicht ziet u de gegevens die op dit moment bij ons zijn geregistreerd. eHerkenning is de digitale toegangssleutel die u steeds vaker nodig zult hebben om bij overheidsinstellingen in te loggen. Om eHerkenning te kunnen gebruiken, hebt u een eHerkenningmiddel met betrouwbaarheidsniveau 2. Gebruikt u bij één van de erkende aanbieders. Zie voor meer informatie op SCHURLURL. KVK-nummer: ODAAdres: ODAAdres: ODAAdres: ODAAdres: ODAAdres: ODAAdres: ODAAdres: ODAAdres: ODAAdres: ODAAdres: ODAAdres: ODAAdres: ODAAdres: Kies, nadat u bent ingelogd, voor Mijn dossier en daarna voor Mijn gegevens wijzigen. Denk hierbij aan het doorgeven van uw rekeningnummer als u subsidies wilt doorgeven of het doorgeven van uw rekeningnummer. Meer informatie Heeft u vragen, kijk dan op SCHURLURL. Of neem telefonisch contact met ons op: ODTelefoonnummer LD (lokaal tarief). De Minister van Landbouw, Natuur en Voedselkwaliteit groet, De Minister van Landbouw, Natuur en Voedselkwaliteit.

Letter 4: 08. Herinnering heffing DGF (FKGL Score: 6th highest)

Original FKGL: 7.2

Original Content:

Betreft: Herinnering heffing Diergezondheidsfonds pluimvee

Geachte heer/mevrouw,

Misschien is het aan uw aandacht ontsnapt, maar u heeft de heffing Diergezondheidsfonds (DGF) pluimvee januari t/m december 2022 nog niet (helemaal) betaald. Daarom sturen we u een herinnering. In deze brief leest u om welk bedrag het gaat. En hoe u dit op tijd naar ons overmaakt.

Het bedrag dat nog openstaat is € 19,62. Wilt u dit voor 26 januari 2024 naar ons overmaken? Gebruik alstublieft de bankgegevens die hieronder staan. In de omschrijving zet u het factuurnummer.

Begunstigde: Rijksdienst voor Ondernemend Nederland IBAN: [IBAN] BIC: [BIC] (voor betalen uit het buitenland) Factuurnummer: [Factuurnummer]

Krijgt u ook andere facturen van ons? DGF heeft een eigen rekening. Het overmaken van deze heffing doet u op bovenstaand DGF-rekeningnummer. Heeft u inmiddels al betaald? Dan hoeft u niks te doen.

Meer informatie U kunt de originele factuur bekijken op [URL] onder Documenten. Heeft u vragen? Kijk op [URL] of bel ons via: [Telefoonnummer] (lokaal tarief).

Met vriendelijke groet,

De Minister van Landbouw, Natuur en Voedselkwaliteit

Rewritten FKGL: 10.3

Rewritten Content:

We hebben de heffing Diergezondheidsfonds (DGF) Herinnering heffing Diergezondheidsfonds (DGF) Herinnering heffing Diergezondheidsfonds (DGF) Herinnering heffing Diergezondheidsfonds (DGF) Herinnering heffing Diergezondheidsfonds (DGF) Herinnering heffing Diergezondheidsfonds (DGF) Herinnering heffing Diergezondheidsfonds (DGF) We sturen u een herinnering. In deze brief leest u dat bedrag het gaat. Hoe u dit op tijd naar ons overmaakt. Het bedrag dat nog openstaat is ICE 19,62. Gebruik alstublieft de bankgegevens die hieronder staan. In de omschrijving zet u het factuurnummer. Gebruikt u ook andere facturen van ons? Het overmaken van deze heffing doet u op bovenstaand DGF-rekeningnummer. Als u nu al betaalt, hoeft u niets te doen. Meer informatie U kunt de originele factuur bekijken op LDURL onder Documenten. Denk op ODURLIND of bel ons via: ODTelefoonnummer. De Minister van Landbouw, Natuur en Voedselkwaliteit groet, De Minister van Landbouw, Natuur en Voedselkwaliteit.

Letter 5: 80. Akkoord overdracht voucher SABE (FKGL Score: 9th highest)

Original FKGL: 6.2

Original Content:

Betreft: Overdracht SABE Advies over duurzame landbouw voor agrariërs (2022)

Geachte heer/mevrouw [Naam],

Op 17 mei 2022 heb ik uw overdracht ontvangen voor SABE1A van de regeling SABE Advies over duurzame landbouw voor agrariërs (2022).

Overdracht Ik bevestig de overdracht van uw voucher en heb deze geregistreerd. De voucher is overgedragen aan: Naam adviseur: [Naam adviseur]

Inhoudelijke wijziging Wilt u na overdracht nog een inhoudelijke wijziging doorgeven? Ga dan naar Mijn zaken en kies de zaak die u wilt veranderen. U kunt daarna bij Optionele acties een inhoudelijke wijziging doorgeven.

Wilt u uw zaak bekijken of wijzigingen doorgeven?

Ga naar [URL] Klik bij 'Direct regelen' op Beheren Log in met e-Herkenning. U kunt nu uw zaken inzien en beheren. Meer informatie Heeft u vragen? Kijk dan op [URL]. Of neem telefonisch contact op: [Telefoonnummer] (lokaal tarief).

Met vriendelijke groet,

Teammanager Regelings Overstijgende Controles

Rewritten FKGL: 9.9

Rewritten Content:

Op 17 mei 2022 heb ik uw overdracht ontvangen voor SABE Advies over duurzame landbouw voor agrariërs (2022). Ik bevestig de overdracht van uw voucher en heb deze geregistreerd. Wilt u na overdracht nog een inhoudelijke wijziging doorgeven? Ga dan naar Mijn zaken en kies de zaak die u wilt veranderen. U kunt daarna bij Optionele acties een inhoudelijke wijziging doorgeven. Wilt u uw zaak bekijken of wijzigingen doorgeven? 1. Ga naar ODURL 2. Klik bij 'Direct regelen' op Beheren 3. Log in met e-Herkenning. U kunt nu uw zaken

Wijzigingen: U voert het project uit conform de door u ingediende opdrachtbevestiging. Bij wijziging van de start- of einddatum, de leverancier of de omvang van de investering dient u vooraf een gemotiveerd verzoek in.

Gecombineerde opgave: Gedurende de gehele projectduur voldoet u aan de wettelijke verplichting voor het invullen van de gecombineerde opgave en (indien op u van toepassing) de CO₂-emissieaangifte.

Vaststelling: Uiterlijk 13 weken na afloop van het project dient u een aanvraag tot vaststelling van de subsidie in. De vaststellingsaanvraag gaat vergezeld van een kort inhoudelijk verslag. Als uw subsidie meer dan €125.000,- bedraagt, dient u tevens een controleverklaring in (overeenkomstig het Controleprotocol in bijlage 1.3 van de Regeling nationale EZK- en LNV-subsidie).

Ik maak u erop attent dat uw project in de steekproef kan vallen. In dat geval kan ik u verzoeken informatie aan te leveren over de inhoudelijke en financiële afhandeling van uw project. Als u zich niet aan de verplichtingen houdt, kan ik de subsidie geheel of gedeeltelijk intrekken en terugvorderen, inclusief wettelijke rente.

Bijlagen In bijlage 1 staat het betaalschema.

Meer informatie over investeringen in energie in de glastuinbouw kunt u vinden op: [URL].

Als u bezwaar wilt maken Bent u het niet eens met deze beslissing? U kunt binnen zes weken na de datum bovenaan deze brief bezwaar maken. Ga naar [URL] om uw bezwaar digitaal te versturen. Kies voor 'Uw bezwaar regelen via eLoket'. Wilt u liever schriftelijk bezwaar maken? Stuur uw bezwaarschrift dan naar Rijksdienst voor Ondernemend Nederland, afdeling Juridische Zaken, [Adres]. Noem hierin het referentienummer en de datum van de beslissing waartegen u bezwaar maakt. U vindt de referentie in de rechter kantlijn van deze brief.

Wilt u uw zaak bekijken of wijzigingen doorgeven? 1. Ga naar [URL]. 2. Klik bij 'Direct regelen' op Beheren. 3. Log in met e-Herkenning. U kunt nu uw zaken inzien en beheren.

Meer informatie Heeft u vragen, kijk dan op [URL]. Of neem telefonisch contact op: [Telefoonnummer] (lokaal tarief).

Met vriendelijke groet,

De Minister van Landbouw, Natuur en Voedselkwaliteit

Rewritten FKGL: 7.1

Rewritten Content:

Op 19 april 2022 heb ik uw aanvraag ontvangen voor EG-22- 2022. Ik heb het besluit genomen om dit verzoek toe te passen. De subsidieverlening bedraagt maximaal TSBedrag. Ik heb ambtshalve besloten voorschotten te verstrekken tot 90% van de verleende subsidie. In de bijlage is het betaalschema met overzicht van de voorschotbedragen te vinden. Ik zal met de subsidievaststelling verrekenen. Hieronder vindt u uw zaakgegevens. De aanvraag voor EHG onderdeel LED belichting vindt plaats omdat er een voorlopig meetrapport is ingediend voor de beoordeling van de voorwaarden voor de subsidie. De beoordeling van drie subsidievoorwaarden is: A: het aantal micromol fotonen bij een golflengte van 500-600 nm. We willen u nogmaals attenderen over de volgende afspraken bij het ingediende en verwerkte voorlopige meetrapport voor de EHG subsidie onderdeel (g) LED belichting: 1) Reeds uitbetaalde subsidie uitbetaalde voorschotten moeten dan terugbetaald worden. Verplichtingen Op deze subsidieverlening zijn de bepalingen van de Regeling nationale EZK- en LNV-subsidies en het Kaderbesluit nationale EZK-subsidies. U kunt zowel het Kaderbesluit als de Regeling nationale EZK-subsidies vinden op internet. Bij de zoekfunctie kunt u bij de zoekfunctie de datum van uw brief invult. U krijgt dan de tekst van het besluit of de regeling te zien zoals die voor dit project gelden.

Als uw project langer duurt dan 12 maanden, rapporteert u jaarlijks over de voortgang van de activiteiten. U voert het project uit conform de door u ingediende opdrachtbevestiging. Bij wijziging van de start- of einddatum, de leverancier of de grootte van de investering moet u vooraf een gemotiveerd verzoek in. Na de gehele projectduur voldoet u aan de wettelijke verplichting voor het invullen van de gecombineerde opgave en (indien op u van toepassing) de CO₂-emissieaangifte. Uiterlijk 13 weken na afloop van het project dient u een aanvraag tot vaststelling van de subsidie in. De vaststellingsaanvraag gaat vergezeld van een kort inhoudelijk verslag. Als uw subsidie meer dan €125.000 bedraagt, moet u tevens een controleverklaring in (overeenkomstig het Controleprotocol in bijlage 1.3 van de Regeling nationale EZK-subsidie. Ik maak u erop dat uw project in de steekproef kan vallen. In dat geval kan ik u verzoeken informatie te leveren over de inhoudelijke en financiële afhandeling van uw project. Als u zich niet aan de verplichtingen houdt, kan ik de subsidie geheel of gedeeltelijk intrekken en terugvorderen, inclusief wettelijke rente. Bijlagen In bijlage 1 staat het betaalschema. Meer informatie over investeringen in energie in glastuinbouw kunt u vinden op: [SCHURLURL](#). Als u bezwaar wilt maken Bent u het niet eens met deze beslissing? Om uw bezwaar digitaal te sturen, ga naar [ODURL LD](#) om uw bezwaar digitaal te versturen. Kies voor 'Uw bezwaar regelen via eLoket'. Stuur uw bezwaarschrift dan naar Rijksdienst voor Ondernemend Nederland, afdeling Juridische Zaken. Dit referentienummer en de datum van de beslissing waartegen u bezwaar maakt. U vindt de referentie in de rechter kantlijn van deze brief. Wilt u uw zaak bekijken of wijzigingen doorgeven? 1. Ga naar [ULURL Scots](#). 2. Klik bij 'Direct regelen' op 'Direct regelen'. 3. Log in met e-Herkenning. U kunt nu uw zaken inzien en beheren. Meer informatie Heeft u vragen, kijk dan op [SCHURLURL](#). Of neem telefonisch contact op: [ODTelefoonnummer](#) (lokaal tarief). De Minister van Landbouw, Natuur en Voedselkwaliteit groet, De Minister van Landbouw, Natuur en Voedselkwaliteit.

Letter 7: 37. Aanvullende gegevens landbouwer (Rank 2: Closest to Mean FKGL)

Original FKGL: 6.1

Original Content:

Betreft: Aanvullende gegevens landbouwer 2023

Geachte heer/mevrouw,

Wij nodigen u uit om aanvullende gegevens door te geven. Hiermee controleren wij of u zich aan de regels van het mestbeleid heeft gehouden. In deze brief leest u wanneer wij de gegevens van u nodig hebben en hoe u deze doorgeeft.

Uiterlijk 31 januari doorgeven Geef uw aanvullende gegevens uiterlijk 31 januari 2024 door. Het formulier voor 2023 staat voor u klaar op [mijn.rvo.nl](#). Ga naar Registratie en melding doorgeven [; Aanvullende gegevens landbouwer](#). Voor het inloggen heeft u uw DigiD of eHerkenning nodig. Heeft u deze nog niet? Vraag deze dan op tijd aan.

Extra belangrijk bij derogatie Heeft u voor 2023 een derogatievergunning aangevraagd? Dan is het extra belangrijk om de gegevens op tijd door te geven. Zo voorkomt u problemen met de vergunningverlening en dat u een volgend jaar mogelijk geen gebruik mag maken van derogatie.

Bedrijfsoverdracht gehad? Geef in ieder geval aanvullende gegevens door voor het bedrijf waarvoor u bent uitgenodigd. Het relatienummer vindt u rechts bovenaan deze brief. Meer weten over aanvullende gegevens doorgeven na een bedrijfsoverdracht? U leest het op [\[URL\]](#) [; Uw landbouwbedrijf](#) [; Aanvullende gegevens landbouwer](#). Kijk bij het kopje Wanneer en hoe doorgeven.

Analysegegevens gebruiken Heeft u een eindvoorraad mest? Dan gebruikt u de best beschikbare gegevens voor het bepalen van de hoeveelheid stikstof en fosfaat. Dat zijn de analyseresultaten van de hele mestvoorraad, of van de in 2023 afgevoerde mest. Heeft u geen analyseresultaten?

Dan gebruikt u de wettelijke normen. Meer hierover leest u op Aanvullende gegevens landbouwer op [URL].

Meer informatie Heeft u vragen? Kijk dan op [URL] voor meer informatie. Chat met ons op [URL]. Of bel ons op: [Telefoonnummer] (lokaal tarief).

Met vriendelijke groet,

De Minister van Landbouw, Natuur en Voedselkwaliteit

Rewritten FKGL: 7.1

Rewritten Content:

Aanvullende gegevens landbouwer 2023 Aanvullende gegevens landbouwer 2023 Aanvullende gegevens landbouwer 2023 Aanvullende gegevens landbouwer 2023 Aanvullende gegevens landbouwer 2023 Aanvullende gegevens landbouwer 2023 Aanvullende gegevens landbouwer 2023 We controleren of u zich aan de regels van het mestbeleid heeft gehouden. In deze brief leest u wanneer wij de gegevens van u nodig hebben en hoe u deze doorgeeft. Geef uw aanvullende gegevens uiterlijk 31 januari 2024 door. Het formulier voor 2023 staat voor u klaar op mijn.rvo.nl. Ga naar Registratie en melding doorgeven TS Aanvullende gegevens landbouwer. Om te kunnen inloggen heb je een DigiD of eHerkenning nodig. Vraag deze dan op tijd aan. Bij derogatie is het belangrijk om de gegevens op tijd door te geven. Als u een volgend jaar geen gebruik mag maken van derogatie, mag u in ieder geval geen gebruik maken van derogatie. Het relatienummer vindt u rechts bovenaan deze brief. Het landbouwbedrijf IND Aanvullende gegevens landbouwer. Denk bij het kopje Wanneer en hoe doorgeven. Om de hoeveelheid stikstof en fosfaat te bepalen, gebruikt u de best beschikbare gegevens voor het bepalen van hoeveel stikstof en fosfaat. Die analyseresultaten van de hele mestvoorraad, of van de in 2023 afgevoerde mest. Als u geen analyseresultaten gebruikt, gebruikt u de wettelijke normen. Meer daarover leest u op Aanvullende gegevens landbouwer op ODURLIND. Denk dan op STURLFOR voor meer informatie. We zijn met ons op SCHURLURL. Of bel ons op: SCHIPTelefoonnummer (lokaal tarief). De Minister van Landbouw, Natuur en Voedselkwaliteit groet, De Minister van Landbouw, Natuur en Voedselkwaliteit.

Letter 8: 38. Aanvullende gegevens landbouwer mineralenconcentraat (Rank 3: Closest to Mean FKGL)

Original FKGL: 6.2

Original Content:

Betreft: Aanvullende gegevens landbouwer 2023

Geachte heer/mevrouw,

Wij nodigen u uit om aanvullende gegevens door te geven. Hiermee controleren wij of u zich aan de regels van het mestbeleid heeft gehouden. In deze brief leest u wanneer wij de gegevens van u nodig hebben en hoe u deze doorgeeft.

Uiterlijk 31 januari doorgeven Geef uw aanvullende gegevens uiterlijk 31 januari 2024 door. Het formulier voor 2023 staat voor u klaar op mijn.rvo.nl. Ga naar Registratie en melding doorgeven j Aanvullende gegevens landbouwer. Voor het inloggen heeft u uw DigiD of eHerkenning nodig. Heeft u deze nog niet? Vraag deze dan op tijd aan.

Extra belangrijk bij derogatie Heeft u voor 2023 een derogatievergunning aangevraagd? Dan is het extra belangrijk om de gegevens op tijd door te geven. Zo voorkomt u problemen met de vergunningverlening en dat u een volgend jaar mogelijk geen gebruik mag maken van derogatie.

Bedrijfsoverdracht gehad? Geef in ieder geval aanvullende gegevens door voor het bedrijf waarvoor u bent uitgenodigd. Het relatienummer vindt u rechts bovenaan deze brief. Meer weten over aanvullende gegevens doorgeven na een bedrijfsoverdracht? U leest het op [URL] ; Uw landbouwbedrijf ; Aanvullende gegevens landbouwer. Kijk bij het kopje Wanneer en hoe doorgeven.

Mineralenconcentraat doorgeven Had u op 31 december 2023 een eindvoorraad mineralenconcentraat? Deze geeft u aan ons door met mestcode 120. De hoeveelheid stikstof en fosfaat berekent u met de analysegegevens. Heeft u deze niet? Dan mag u rekenen met de wettelijke normen (6,4 kg stikstof per ton en 0,3 kg fosfaat per ton).

Analysegegevens gebruiken Heeft u een eindvoorraad mest? Dan gebruikt u de best beschikbare gegevens voor het bepalen van de hoeveelheid stikstof en fosfaat. Dat zijn de analyseresultaten van de hele mestvoorraad, of van de in 2023 afgevoerde mest. Heeft u geen analyseresultaten? Dan gebruikt u de wettelijke normen. Meer hierover leest u op Aanvullende gegevens landbouwer op [URL].

Meer informatie Heeft u vragen? Kijk dan op [URL] voor meer informatie. Chat met ons op [URL]. Of bel ons op: [Telefoonnummer] (lokaal tarief).

Met vriendelijke groet,

De Minister van Landbouw, Natuur en Voedselkwaliteit

Rewritten FKGL: 7.1

Rewritten Content:

Aanvullende gegevens landbouwer 2023 Aanvullende gegevens landbouwer 2023 Aanvullende gegevens landbouwer 2023 Aanvullende gegevens landbouwer 2023 Aanvullende gegevens landbouwer 2023 Aanvullende gegevens landbouwer 2023 Aanvullende gegevens landbouwer 2023 We controleren of u zich aan de regels van het mestbeleid heeft gehouden. In deze brief leest u wanneer wij de gegevens van u nodig hebben en hoe u deze doorgeeft. Geef uw aanvullende gegevens uiterlijk 31 januari 2024 door. Het formulier voor 2023 staat voor u klaar op mijn.rvo.nl. Ga naar Registratie en melding doorgeven TS Aanvullende gegevens landbouwer. Om te kunnen inloggen heb je een DigiD of eHerkenning nodig. Vraag deze dan op tijd aan. Bij derogatie is het belangrijk om de gegevens op tijd door te geven. Als u een volgend jaar geen gebruik mag maken van derogatie, mag u in ieder geval geen gebruik maken van derogatie. Het relatienummer vindt u rechts bovenaan deze brief. Het landbouwbedrijf IND Aanvullende gegevens landbouwer. Mineralenconcentraat doorgeven Had u op 31 december 2023 een eindvoorraad mineralenconcentraat? De hoeveelheid stikstof en fosfaat berekent u met de analysegegevens. Kwantumregels (6,4 kg stikstof per ton en 0,3 kg fosfaat per ton). Om de hoeveelheid stikstof en fosfaat te bepalen, gebruikt u de best beschikbare gegevens voor het bepalen van hoeveel stikstof en fosfaat. Die analyseresultaten van de hele mestvoorraad, of van de in 2023 afgevoerde mest. Als u geen analyseresultaten gebruikt, gebruikt u de wettelijke normen. Meer daarover leest u op Aanvullende gegevens landbouwer op ODURLIND. Denk dan op STURLFOR voor meer informatie. We zijn met ons op SCHURLURL. Of bel ons op: SCHIPTelefoonnummer (lokaal tarief). De Minister van Landbouw, Natuur en Voedselkwaliteit groet, De Minister van Landbouw, Natuur en Voedselkwaliteit.

A.3 Letters with Biggest FKGL Score Improvements

Letter 9: 05. Factuur Binnenvisserij (Rank 1: Greatest Improvement in FKGL)

Original FKGL: 15.4

Original Content:

FACTUUR

Geachte heer/mevrouw,

Voor de beoordeling van uw verzoek met bovengenoemd verzoeknummer, brengen wij u in-gevolge de Visserijwet een bedrag van € 30,00 in rekening.

Ik verzoek u vriendelijk het bovenstaande bedrag binnen 30 dagen te voldoen op de bankrekening van ING Bank, rekeningnummer [IBAN], ten name van RIJKSDIENST VOOR ONDERNEMEND NEDERLAND onder vermelding van het bovenstaande verzoeknummer.

Het verzoek wordt afgehandeld zodra de betaling binnen is.

Bij correspondentie dient u het betreffende verzoeknummer te vermelden.

Rewritten FKGL: 9.1

Rewritten Content:

Om uw verzoek met bovengenoemd verzoeknummer te beoordelen, brengen wij u in-gevolge de Visserijwet een bedrag van € 30,00 in rekening. Ik verzoek u vriendelijk het bovenstaande bedrag binnen 30 dagen te voldoen op de bankrekening van ING Bank. Bij correspondentie moet u het verzoeknummer vermelden.

Letter 10: 30. Factuur Grondkamer Zuidwest (Rank 2: Greatest Improvement in FKGL)

Original FKGL: 15.5

Original Content:

Factuur Grondkamer Zuidwest

U ontvangt deze factuur van de Grondkamer Zuidwest via RVO omdat RVO de financiële administratie van de Grondkamers verzorgt.

Namens de Grondkamer verzoeken wij u het totaalbedrag over te maken op ING bank, IBAN [IBAN], ten name van Rijksdienst voor Ondernemend Nederland, BIC [BIC], onder vermelding van het factuurnummer [Factuurnummer] en relatienummer [Relatienummer].

Heeft u vragen over deze factuur dan kunt u contact opnemen met de afdeling klantcontact, telefoon [Telefoonnummer] of via [URL].

Rewritten FKGL: 11.5

Rewritten Content:

RVO stuurt deze factuur van de Grondkamer Zuidwest via RVO omdat RVO de financiële administratie van de Grondkamers verzorgt. Namens de Grondkamer verzoeken wij u het totaalbedrag over te maken op de afdeling klantcontact, telefoon LDRelatienummer.

Letter 11: 65. UPnl Afwijzen aanvraag inhoudelijk SPUK (Rank 3: Greatest Improvement in FKGL)

Original FKGL: 9.4

Original Content:

Betreft: Besluit tot afwijzing subsidieaanvraag

Geachte heer [Naam],

Op 23 mei 2022 heb ik uw aanvraag ontvangen voor [Projectnaam] van de regeling Specifieke uitkering impulsaanpak winkelgebieden (Spuk iw 2022).

Beslissing Naar aanleiding van de beoordeling van uw aanvraag, heb ik het besluit genomen om dit verzoek af te wijzen.

Toelichting op de afwijzing Op grond van artikel 8, onderdeel b van de Regeling specifieke uitkering impulsaanpak winkelgebieden besluit de minister afwijzend op een aanvraag indien aan het project na toepassing van artikel 10, eerste lid, op één van de onderdelen minder dan 5,5 punten zijn toegekend. U scoorde op rangschikkingscriterium 'Kosteneffectiviteit' minder dan 5,5 punten. De reden dat u op dit criterium een onvoldoende scoort is de volgende:

- In het financiële indieningsspreadsheet voert u kosten op voor 'subsidie aan derden'. De commissie oordeelt dat de berekende private onrendabele top en daarmee dit bedrag niet realistisch is. Ten eerste omdat voor alle categorieën (prijsklassen) woningen een onrendabel deel opgevoerd wordt. Voor de hogere prijsklassen is dit zeer onwaarschijnlijk. Daarnaast is bij de berekening van de onrendabele top geen rekening gehouden met het fiscale voordeel van 21

Vanuit de beoordeling willen wij u verder het volgende meegeven:

- De commissie waardeert de ingreep om het winkelareaal fors te reduceren, maar raadt aan meer aandacht te besteden aan hoe een toekomstbestendige winkelstructuur te realiseren is. - In het financiële indieningsspreadsheet voert u plankosten op waarvan de commissie over de hoogte geen oordeel kunnen vellen ten aanzien van plausibiliteit. Dit heeft niet bijgedragen aan de onvoldoende score op rangschikkingscriterium 'Kosteneffectiviteit', maar indien u een verbeterde aanvraag in wilt dienen tijdens het tweede indieningstijdvak, dan is het advies om deze kostenpost beter te onderbouwen. De specifieke uitkering wordt namelijk na afronding vastgesteld op basis van de daadwerkelijk gemaakte kosten. Indien deze lager blijken dan de niet gevalideerde kosten zoals opgevoerd in het spreadsheet, dan zou dit kunnen leiden tot het lager vaststellen en deels terugvorderen van een eventuele uitkering.

Als u bezwaar wilt maken Bent u het niet eens met deze beslissing? U kunt binnen zes weken na de datum bovenaan deze brief bezwaar maken. Ga naar [URL] om uw bezwaar digitaal te versturen. Kies voor 'Uw bezwaar regelen via eLoket'.

Wilt u liever schriftelijk bezwaar maken? Stuur uw bezwaarschrift dan naar Rijksdienst voor Ondernemend Nederland, afdeling Juridische Zaken, [Adres]. Noem hierin het referentienummer en de datum van de beslissing waartegen u bezwaar maakt. U vindt de referentie in de rechter kantlijn van deze brief.

Wilt u uw zaak bekijken? 1. Ga naar [URL]. 2. Klik bij 'Direct regelen' op Beheren. 3. Log in met e-Herkenning. U kunt nu uw zaken inzien en beheren.

Meer informatie Heeft u vragen, kijk dan op [URL]. Of neem telefonisch contact op: 088 042 42 42 (lokaal tarief).

Met vriendelijke groet,

De Minister van Economische Zaken en Klimaat

Rewritten FKGL: 6.2

Rewritten Content:

Op 23 mei 2022 heb ik uw aanvraag ontvangen voor SCHuk iw 2022. Ik heb het besluit genomen om dit verzoek af te wijzen. Toelichting op de afwijzing Op grond van artikel 10, eerste lid, op één van de onderdelen minder dan 5 punten toegekend. U scoorde op rangschikkingscriterium ODKosteneffectiviteit U scoorde minder dan 5 punten. In het financiële indieningsspreadsheet gaat u op dit criterium een onvoldoende scoort. De commissie oordeelt dat de berekende private onrendabele top en daarmee dit bedrag niet realistisch is. Dit is voor alle categorieën (prijsklassen) woningen een onrendabel deel. Dit is waarschijnlijk voor de hogere prijzenklassen.

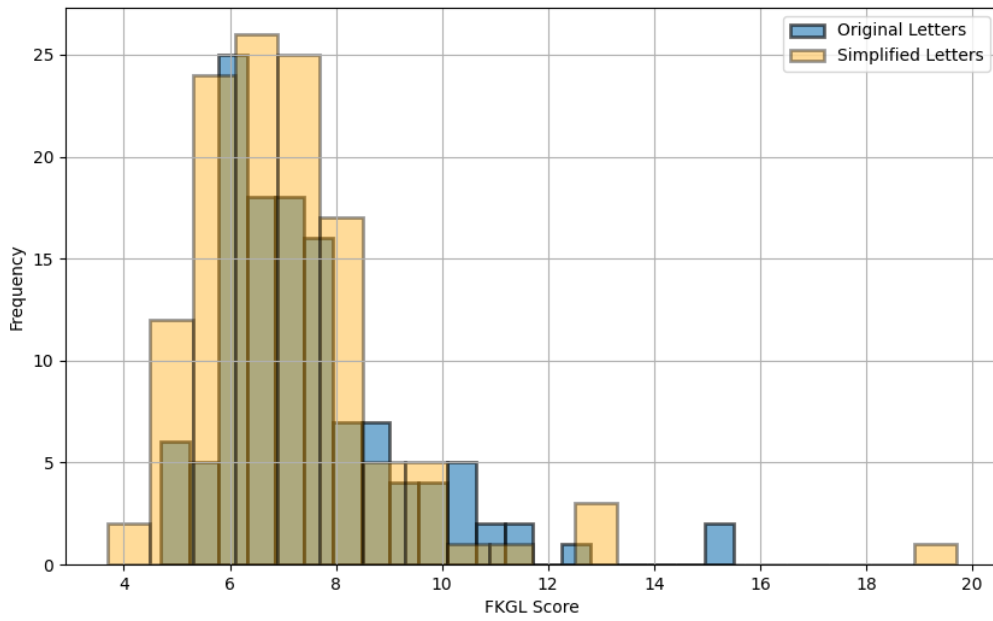


Figure 8: Combined Histogram of FKGL Scores for Original and Simplified Letters

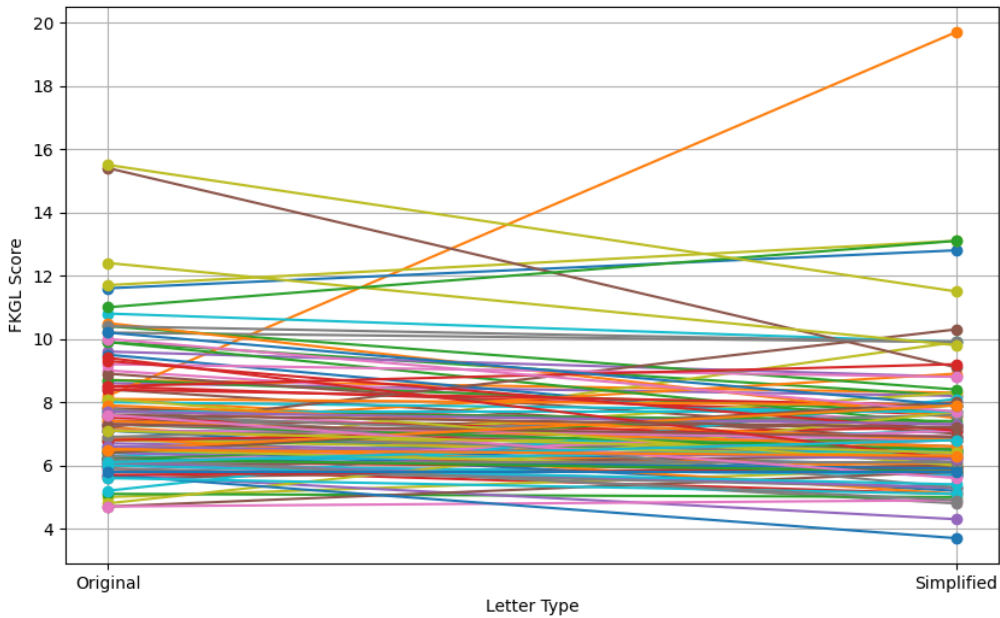


Figure 9: Paired FKGL Scores for Original and Simplified Letters