



# Abstractive Text Summarization for Icelandic

Þór Sverrisson



Faculty of Industrial Engineering, Mechanical Engineering and  
Computer Science  
University of Iceland  
2023



# ABSTRACTIVE TEXT SUMMARIZATION FOR ICELANDIC

Pór Sverrisson

60 ECTS thesis submitted in partial fulfillment of a  
*Magister Scientiarum* degree in Computer Science

Advisor

Hafsteinn Einarsson

Faculty Representative

Steinn Guðmundsson

M.Sc. Committee

Vésteinn Snæbjarnarson

Faculty of Industrial Engineering, Mechanical Engineering and Computer  
Science

School of Engineering and Natural Sciences

University of Iceland

Reykjavik, January 2023

Abstractive Text Summarization for Icelandic  
Abstractive Text Summarization for Icelandic  
60 ECTS thesis submitted in partial fulfillment of a M.Sc. degree in Computer Science

Copyright © 2023 Þór Sverrisson  
All rights reserved

Faculty of Industrial Engineering, Mechanical Engineering and Computer Science  
School of Engineering and Natural Sciences  
University of Iceland  
Dunhagi 5  
107, Reykjavik, Reykjavik  
Iceland

Telephone: 525 4000

Bibliographic information:

Þór Sverrisson, 2023, Abstractive Text Summarization for Icelandic, M.Sc. thesis, Faculty of Industrial Engineering, Mechanical Engineering and Computer Science, University of Iceland.

Printing: Háskólaprent, Fálkagata 2, 107 Reykjavík  
Reykjavik, Iceland, January 2023



# Abstract

In this study, we train and evaluate state-of-the-art Transformer models for abstractive text summarization in Icelandic. We introduce the first publicly available abstractive summarization dataset for Icelandic, RÚV Radio News (RRN), which we use for training and evaluating our models. Our findings reveal that when fine-tuned for a low-resource summarization setting, the multilingual mT5 model outperforms a similar capacity PEGASUS model that is pre-trained exclusively on Icelandic text. Additionally, we discover that data augmentation, using machine-translated text, can result in ungrammatical output text.

# Útdráttur

Verkefnið snýr að því að þjálfa og meta gæði state-of-the-art Transformer líkana í abstrakt samantektargerð fyrir íslensku. Við kynnum til sögunnar fyrsta abstrakt samantektargerðargagnasafnið, RÚV Radio News (RRN), unnið upp úr hádegisfréttum Ríkisútvarpsins. Við berum saman fjöltyngda mT5 málalíkan við sambærilegt PEGASUS líkan, forþjálfað einungis á íslenskum texta. Okkar niðurstöður sýna að mT5 stendur sig almennt betur þegar líkönin eru finþjálfuð fyrir samantektargerð. Ennfremur komumst við að því að þegar við finþjálfum á vélþýddum samantektargerðargagnasöfnum fáum við gjarnan málfræðilega verri úttakstexta.



# Contents

List of Figures	ix
List of Tables	xi
Acknowledgments	xiii
1. Introduction	1
2. Literature Review	3
3. Methods	7
3.1. Data . . . . .	7
3.1.1. Pre-training Corpus . . . . .	7
3.1.2. Fine-tuning Corpora . . . . .	8
3.1.3. Data Preprocessing . . . . .	8
3.2. Models . . . . .	9
3.2.1. Gullfaxi . . . . .	9
3.2.2. mT5 . . . . .	10
3.3. Training . . . . .	10
3.3.1. Pre-training Objective . . . . .	10
3.3.2. Vocabulary . . . . .	11
3.3.3. Hyperparameter configuration . . . . .	12
3.3.4. Downstream tasks . . . . .	12
3.4. Performance . . . . .	13
3.4.1. ROUGE . . . . .	13
3.4.2. Human Evaluation . . . . .	14
4. Results	15
4.1. Model Comparison . . . . .	15
4.2. Low-resource Fine-tuning . . . . .	16
4.3. Fine-tuning Data Augmentation . . . . .	18
4.4. Human Evaluation . . . . .	19
5. Discussion	21
5.1. Abstractive Summarization Data . . . . .	21
5.2. Pre-training of Gullfaxi . . . . .	22



## Contents

5.3. Low-resource Languages Benefit from Multilingual Pre-training . . . .	23
5.4. The Value of Adding Fine-tuning Examples . . . . .	24
5.5. Limitations of the Approach . . . . .	24
5.6. Machine Translations for Data Augmentation . . . . .	25
5.7. Evaluating Abstractive Summaries . . . . .	25
5.8. Conclusion . . . . .	26
Bibliography	29
A. Translated Datasets' Examples	35
A.1. XSum . . . . .	35
A.2. CNN/DailyMail . . . . .	40
B. Hyperparameters	45
C. Example Model Outputs	47
D. Human Evaluation Examples	67

# List of Figures

- 3.1. Gap Sentence Generation. . . . . 11
  
- 4.1. Pre-training validation loss comparison of the Gullfaxi models. . . . . 16
- 4.2. Low-resource fine-tuning . . . . . 17



# List of Tables

- 3.1. Overview of the datasets used in this study. . . . . 7
  
- 4.1. Model comparison when fine-tuned on RRN . . . . . 15
- 4.2. Fine-tuning data augmentation using the Icelandic translations of XSum and CNN/DailyMail. . . . . 18
- 4.3. Performance on the Icelandic translations of XSum and CNN/DailyMail. 19
- 4.4. The human evaluation scores for the Main → Intro task. . . . . 19
- 4.5. Ungrammatical output of Gullfaxi when fine-tuned exclusively on the Icelandic translations of CNN/DailyMail. . . . . 20



# Acknowledgments

Many thanks to Prof. Dr. - Ing. Morris Riedel and the team at Forschungszentrum Jülich for providing access to their supercomputing center.

Thanks should also go to the Icelandic National Broadcasting Service (RÚV) for providing access to their database.

A special thanks to my instructor, Hafsteinn Einarsson, for his invaluable assistance on this project.

Lastly, I would like to thank my family and friends for their continuous support.



# 1. Introduction

Text summarization is the task of producing a shortened version of a text document that preserves the key information of the original text. It is used for a variety of applications, ranging from summarizing medical documents, scientific papers, or legal text to blog posts and news articles. With the increasing amount of available information in our everyday life, there is a high demand for well-written summaries that are coherent and flow naturally from one sentence to the next. Manual text summarization can be a challenging and time-consuming task and the interest in automatic summarization has been growing in recent years.

Automatic text summarization methods can be broadly categorized into two approaches (Tas and Kiyani, 2017). An extractive summarization method composes the summary out of copies of important sections from the original text source. An abstractive summarization method on the other hand rephrases and shortens the original text similar to how a human would approach the task. An abstractive summary should be linguistically fluent and can include novel words whereas an extractive summary is constrained by the original text.

Recently, there have been great advances in the field of natural language processing (NLP) using machine learning methods. This advancement can largely be attributed to the rise of Transformer models (Vaswani et al., 2017) which have become the foundation of such tasks. Transformers can be used for a variety of different problems and have proven successful for automatic text generation tasks such as summarization.

In NLP studies on Transformers, high-resource languages such as English are commonly used with their readily available and diverse text corpora. Machine learning models derived from such studies often rely on a large amount of text data and are computationally expensive to train. Applying these state-of-the-art methods directly to low-resource languages can be hard as they do not always perform well in data-scarce settings (Zoph et al., 2016).

The requirement for extensive text and computational resources prompts us to question whether state-of-the-art Transformer models can be adapted for use in low-resource settings. Icelandic is an interesting case study, a language that has a rich literary history but is spoken by less than 400 thousand people. With the increasing use of technology in everyday life, one step forward to preserving languages such as



## 1. Introduction

Icelandic is by developing NLP solutions.

The main challenge in developing machine learning solutions for low-resource languages is the lack of data. Large models usually call for a large volume of data and specific tasks require specific datasets. Training a model for a particular domain such as news requires a training dataset sourced from news articles. Furthermore, for summarization, the dataset will have to include pairs of text input along with its summary. State-of-the-art Transformer models often use terabytes of training data and large computational resources are needed to conduct these experiments (Hoffmann et al., 2022; Raffel et al., 2019).

One way forward for low-resource languages is to apply transfer learning methods (Zhuang et al., 2021). Such methods have gained popularity when it comes to training deep neural networks with limited data. Transfer learning makes use of a pre-trained model that has been trained on a related task with the intent to transfer knowledge to improve the learning of a new task. We can assume that a network that has already learned patterns and forms of language needs fewer resources to learn a specific NLP task like summarization.

Another way forward for dealing with low-resource problems is to apply data augmentation techniques (Tanner and Wong, 1987) where the amount and diversity of the training data are increased. Such augmentations are achieved by algorithms that generate new synthetic training examples from existing data. In the case of dealing with low-resource settings, we can for example use machine-translated data examples (Sennrich et al., 2016).

This study aims to investigate the potential of transfer learning and data augmentation techniques in achieving high-quality abstractive summarization for Icelandic, a language with limited summarization resources. We compare the use of a pre-trained multilingual model to one that is pre-trained from scratch on Icelandic text. Specifically, we employ the self-supervised pre-training objective proposed by PEGASUS (Zhang et al., 2020) to leverage diverse text data in an effort to compensate for the scarcity of summarization data. The study seeks to determine if these methods can effectively enhance summarization in languages such as Icelandic, where both pre-training and fine-tuning data are highly limited.

Our work primarily evaluates the value of existing models and methods in a low-resource setting. We explore the limits of our approach and the problems that arise when working with morphologically rich languages such as Icelandic. We introduce abstractive summarization models for Icelandic and compare their performance. With the study, we release the first publicly available abstractive summarization dataset for Icelandic, RÚV Radio News (RRN), processed from news stories from The Icelandic National Broadcasting Service (RÚV).

## 2. Literature Review

Up until a few years ago, abstractive summarization methods were considered challenging to build as they require not only extracting the key information from a text but also having the linguistic ability to rewrite and rephrase the content. Although there have previously been studies achieving promising performance using recurrent neural networks and attention-based models (Chopra et al., 2016; Nallapati et al., 2016; Rush et al., 2015), their focus was rather towards sentence level summarization. However, the introduction of the Transformer architecture (Vaswani et al., 2017) has brought forth models that can efficiently work with larger text segments and attend to contextual information in the input.

### The Transformer Model

The Transformer architecture is currently the model of choice for NLP tasks. Exploiting the attention mechanism along with parallel computing, it is ideal for processing sequential data such as natural language text. Compared to sequence models like recurrent neural networks, the Transformer can process the entire input sequence simultaneously rather than sequentially, allowing for parallel processing and effective usage of GPUs.

Transformer models are usually composed of an encoder and/or decoder modules. The encoder maps an input sequence to a sequence of embeddings which are received by the decoder and converted to an output sequence. Depending on the task, Transformer-based models can be limited to an encoder-only or decoder-only architecture. The encoder can process the input in parallel whereas the decoder usually generates the output sequentially. Both the encoder and the decoder consist of multiple layers with trainable parameters. The total number of parameters determines the size of the model.

The computations within the Transformer are based on a mechanism called attention but that mechanism was first introduced by Bahdanau et al. (2015). With attention, the model can learn which parts of the input sequence are relevant to each other for a given task. In contrast to earlier approaches, attention allows the model to

## 2. Literature Review

capture long-range context, which can be essential when dealing with text data. For example, for language comprehension, it is critical to understand the semantic relations between words in a sentence. Incorporating context can be contrasted with older bag-of-word approaches where words were viewed as independent entities unrelated to each other.

### Context Length for Transformers

Transformers avoid the sequential dependency of recurrent neural networks and are able to comprehend longer sequences of text (Al-Rfou et al., 2019). However, the sequence length they can process is limited. For example, the popular BERT model can process sequences with up to 512 tokens (Devlin et al., 2019) and larger models such as GPT-3 can process up to 2048 tokens (Brown et al., 2020). The sequence length in Transformers is inherently limited by the quadratic cost of the attention mechanism, which limits the application of standard Transformer models to documents with at most a few thousand words.

Liu et al. (2018) demonstrated the use of Transformers for multi-document summarization by using an extraction method followed by an abstractive summarization Transformer. Furthermore, methods have been explored with the aim of reducing the computational cost of attention to processing longer sequences. Models such as the Longformer (Beltagy et al., 2020) and BigBird (Zaheer et al., 2020) scale linearly with the sequence length, reducing the resource usage of the standard self-attention mechanism thus allowing for even longer sequences to be processed. However, methods that reduce the computational cost of attention do so at a cost, that is, the attention is restricted, i.e., not all-to-all.

## Transfer learning

Transfer learning is the process of training a model in one task and transferring the knowledge acquired to improve the learning of another task. The application domain of transfer learning is wide but such methods are frequently used for machine learning problems having to do with image or language data. For text data, the intuition is that networks that have been pre-trained on large corpora learn to recognize common patterns and linguistic features and therefore adapt faster to specific language tasks when fine-tuned. Critically, it takes a much shorter time and requires less training data to fine-tune a pre-trained language model than to train one from scratch and the quality of the result will also be better.

## Self-supervised Pre-training of Language Models

Self-supervised learning methods allow training language models on unlabeled text data. This approach has greatly impacted NLP by enabling models to be trained on large amounts of data. Typically, models learn unobserved parts of the input from observed parts. One example of a self-supervised representation learning goal in NLP is Masked Language Modeling (MLM) where some tokens are masked from the text input and the model must predict the missing tokens. Examples of self-supervised pre-training include training word embedding models like Word2Vec (Mikolov et al., 2013) and Transformer language models like GPT (Radford et al., 2018) and BERT (Devlin et al., 2019).

### Pre-training for Summarization

Text generation tasks, like summarization, require Transformer models to have a decoder block to produce text. The sequence-to-sequence (Sutskever et al., 2014) framework using encoder-decoder architectures is a popular approach to abstractive summarization tasks. Pre-training such models for summarization often entails a self-supervised pre-training objective that hides a portion of the input sequence and asks the model to fill the gaps. Song et al. (2019) proposed reconstructing a randomly selected fragment from a sentence given the remaining part of the sentence, achieving quality results when fine-tuned for a number of low-resource summarization tasks.

Numerous NLP problems can be treated as text-to-text tasks as demonstrated by the T5 model (Raffel et al., 2019). This generalization allows for the pre-training of a single encoder-decoder Transformer on a diverse set of tasks. T5 suggested that larger models tended to perform better when using a large volume of pre-training data. Dong et al. (2019) proposed unified modeling, pre-training on three types of language modeling tasks: unidirectional, bidirectional, and sequence-to-sequence prediction. Their approach uses specific self-attention masks to control what context the prediction conditions on.

BART models (Lewis et al., 2019) are trained to reconstruct a text document that has been corrupted with an arbitrary noising function. For pre-training, they use noising techniques such as replacing spans of text with a mask token or randomly permuting the order of sentences. When applied to a wide range of end tasks, BART has proved to be very effective at tasks such as summarization.

Zhang et al. (2020) proposed that using a pre-training objective that closely resembles the summarization task would result in a model that adapts faster when fine-tuned on a small number of examples. They introduced the PEGASUS model

## 2. Literature Review

that learned from masking multiple sentences from an input document and concatenating them to a summary-like output text. PEGASUS adapted well to unseen summarization datasets with as little as 1k fine-tuning examples.

## Data Augmentation

Data augmentation techniques are used to increase the number of data examples artificially and can be beneficial for dealing with low-resource settings (Fadaee et al., 2017). Data augmentation can be viewed as a type of regularization, preventing the model from overfitting to a small number of training examples. Machine translations can be used for augmentation of the training data and such methods can improve a model’s performance in a low-resource setting (Sennrich et al., 2016).

## Applying NLP to Various Languages

Most NLP studies focus on high-resource languages such as English, leaving the vast majority of languages understudied. Low-resource languages face various problems in developing linguistic corpora regarding data collection as well as annotation and representation (Magueresse et al., 2020). However, with the development of text corpora, for those languages opportunities arise for developing NLP solutions in various languages and evaluating existing models in different settings.

With the growing body of work on transfer learning and data augmentation, promising approaches for low-resource NLP have surfaced. This advancement has given rise to Icelandic studies on high-level NLP tasks, ranging from translation models (Símonarson et al., 2021) and question answering (Snæbjarnarson and Einarsson, 2022b) to part-of-speech tagging and named entity recognition (Snæbjarnarson et al., 2022). Studies on summarization for Icelandic are however limited and have only surveyed extractive methods (Christiansen, 2014; Daðason et al., 2021; Daðason and Loftsson, 2022).

Cross-lingual transfer learning has gained widespread attention in recent years, with multilingual models such as XLM-R (Conneau et al., 2020) and mT5 (Xue et al., 2021) achieving state-of-the-art performance in various NLP tasks. Studies on NLP in Icelandic have demonstrated that using pre-trained multilingual models can bring significant benefits for a wide range of tasks (Snæbjarnarson and Einarsson, 2022a; Snæbjarnarson et al., 2022).

# 3. Methods

## 3.1. Data

An overview of the text corpora used in this study is shown in Table 3.1. The English language corpora are translated to Icelandic using machine translation as described in Section 3.1.3. Examples from the translated datasets can be found in Appendix A.

*Table 3.1: Overview of the datasets used in this study. The language column refers to the original language of the dataset.*

<b>Dataset</b>	<b># Documents</b>	<b>Language</b>	<b>Type</b>
IGC 20.05	5M	is	Generic
IGC-News1 21.05 (2020)	112k	is	Generic
RRN	4k	is	Summarization
XSum	227k	en	Summarization
CNN/DailyMail	311k	en	Summarization

### 3.1.1. Pre-training Corpus

For the pre-training of the Gullfaxi model described in Section 3.2.1, we used the Icelandic Gigaword Corpus (IGC, Steingrímsson et al. (2018)) version 20.05. The corpus consists of a text collection of roughly 5 million documents. The documents come from various categories such as adjudications, parliamentary speeches, news, books, and scientific journals. The texts have been automatically divided into sentences and running words, tagged, and lemmatized. For validation during pre-training, we used the IGC-News1 21.05 dataset which consists of documents from the year 2020. These news articles were not included in the training data.

### 3. Methods

#### 3.1.2. Fine-tuning Corpora

For fine-tuning, we used the following news summarization datasets.

- **RÚV Radio News (RRN)** includes news stories from The Icelandic National Broadcasting Service (RÚV). These stories were sourced from the noon radio news program (Hádegisfréttir) for the purpose of this study and include a title, an intro, the main story, and a summary. This dataset was specifically collected for this study and had not been officially released (see additional information below). The dataset is limited to the years 2021 and 2022 and contains a high number of stories related to COVID-19 and domestic news. The fine-tuning dataset does not overlap with the pre-training data.
- **XSum** (Narayan et al., 2018) is an English language dataset that features BBC articles from 2010 to 2017 covering a wide variety of domains. Each article is accompanied by a professionally written single-sentence abstractive summary.
- **CNN/DailyMail** (Hermann et al., 2015) consists of news stories in English featured on CNN and Daily Mail websites. Although not created for summarization tasks, each article is accompanied by human-written abstractive summary bullets.

#### 3.1.3. Data Preprocessing

##### RÚV Radio News (RRN)

The Icelandic National Broadcasting Service (RÚV) granted access to a database of news stories via a custom interface that was only available on-premises at their headquarters. The stories were manually selected from the database, and only transcripts of radio news from 2021 and 2022 were used. The raw output from this system is referred to as the RÚV radio news transcripts.

The RRN dataset was extracted from the RÚV radio news transcripts and comprises four parts for each story: a title, an intro, the main story, and a summary. However, it is worth noting that some stories may not include all parts. The intro and the summary are often similar as they both provide an overview of the key points of the story, and in some instances, they may be identical. For a given date, the summaries were in a separate document and not linked to a story by any unique identifier. Therefore, we matched the summary with the corresponding story by comparing the number of matching unigrams, selecting the pairing with the highest ROUGE1-

F1 score (see Section 3.4.1 for more information). We reviewed 100 random pairings and found that this approach produced correct pairings in all cases.

Stories that do not fit the summarization task, for example, live broadcasts and weather news were filtered out using a label in the transcript. Additionally, we programmatically removed reporters’ comments, phone numbers, and instructions for the broadcast, which were denoted using specific identifiers.

## English to Icelandic Text Translation

For data augmentation, we automatically translated the two abstractive summarization datasets, XSum and CNN/DailyMail from their original English language to Icelandic. The Icelandic translations were obtained using Facebook’s multilingual model winning submission to the 2021 Conference on Machine Translation (WMT, Tran et al. (2021)). The model is fine-tuned for news domain data, trained using data from eight different languages, and achieves state-of-the-art performance in machine translation.

The translation model is implemented in HuggingFace’s Transformers library and the weights were loaded from the pre-trained `wmt21-dense-24-wi de-en-x` repository. The model is trained to translate English text to different languages by forcing a target language id as the first generated token. We noticed during a manual review of the model output that long text input resulted in lower-quality translations. To improve translation performance we split the texts into sentences and translated each sentence separately. Based on manual inspection, we believe that splitting the sentences resulted in better translation results, but we did not perform any quantitative evaluation to confirm that.

## 3.2. Models

### 3.2.1. Gullfaxi

Gullfaxi is based on the PEGASUS model (Zhang et al., 2020) but we train it on Icelandic text. The base architecture is a standard Transformer encoder-decoder but for the study, we compare two different-sized models. Both models, Gullfaxi<sub>BASE</sub> and Gullfaxi<sub>LARGE</sub> correspond to architectures presented in the PEGASUS study. Gullfaxi<sub>BASE</sub> has 223M trainable parameters compared to 568M parameters for Gullfaxi<sub>LARGE</sub>. Considering the difference in the number of parameters, Gullfaxi<sub>LARGE</sub> requires considerably more computational resources for training.



### 3. Methods

#### 3.2.2. mT5

The mT5 model (Xue et al., 2021) is a multilingual variant of T5 (Raffel et al., 2019) that utilizes the unified text-to-text framework in a multilingual setting. mT5 is a Transformer encoder-decoder model pre-trained on a Common Crawl-based dataset, mC4, which covers 101 languages, including Icelandic. In this study, we use the mT5<sub>BASE</sub> version, which has 580 million parameters. The increase in parameter count compared to Gullfaxi<sub>BASE</sub> is primarily due to the larger vocabulary employed in mT5.

## 3.3. Training

### 3.3.1. Pre-training Objective

Gullfaxi is pre-trained using a self-supervised pre-training objective called gap sentence generation (GSG). GSG is a sequence-to-sequence pre-training objective proposed for the pre-training of the PEGASUS model. GSG entails masking whole sentences from the input document and concatenating them in the order they appear in the original document into an abstractive summary-like output text. The intuition is that a pre-training objective that resembles the summarization task leads to a better starting point for fine-tuning.

The gap sentences are chosen using a heuristic approach based on importance criteria. Ideally, the sentences holding the key information of the text are chosen from the text document but that is not guaranteed by the approach. The importance of a sentence within a document is estimated by obtaining the ROUGE1-F1 score between the gap sentence and the rest of the document. In this study, we compute the ROUGE1-F1 score based on the lemmatized sentences as processed in the IGC.

The reason we computed the ROUGE-F1 score on the lemmatized sentences is because of the peculiarities of the Icelandic language. Words in Icelandic are inflected and for that reason, we end up with words with the same meaning that can have different spellings. Lemmatization allows us to compare the base form of each word, which is effectively the same as how the method was applied for English since it does not have the same degree of inflections as Icelandic does. Computing the ROUGE-F1 score in this manner was thus a necessity and those replicating this work should consider specifically if it is relevant to their language as well. We did not perform an ablation study where the ROUGE-F1 score was computed in a different manner but we expect that a straightforward application could have resulted in a lower summarization performance.

**MASK** Var hún fyrsta þotan sem Íslendingar eignuðust og kom hingað til lands sumarið 1967. Þotan er nú í eigu flutningafyrirtækisins UPS og sinnir fraktflutningum. **MASK** „Þetta var afskaplega góð vél og ég saknaði hennar mjög“ sagði Jóhannes í gær. Hann flaug þotunni m.a. til áfangastaða á Norðurlöndum, í sólarlandaferðir og til Grænlands. Jóhannes hætti störfum árið 1980 og lauk flugstjóraferli sínum á þessari þotutegund. Gamli Gullfaxi hefur ekki komið til Íslands frá því hann var seldur árið 1984.

Þrjátíu og sex ára þota af gerðinni B727-100 hafði viðdvöl á Ke-flavíkurflugvelli í gær en þotan hét upphaflega Gullfaxi og var í eigu Flugfélags Íslands. Jóhannes R. Snorrason, fyrrverandi flugstjóri, sótti þotuna fyrir Flugfélag Íslands til Seattle snemmsumars 1967 og lenti henni á Reykjavíkurflugvelli en hún var smíðuð sérstaklega fyrir FÍ.

Figure 3.1: Example of a gap sentence generation. The highlighted sentences have the highest ROUGE1-F1 score and are used as a summary-like output.

The highest scoring models from the PEGASUS study were obtained by choosing a gap sentence ratio between 15%-45%, varying by task and model. For our study, we mask 20% of the total number of sentences in the original text document. Figure 3.1 shows an example where the highlighted sentences have been extracted from the original input document and replaced by a mask token.

### 3.3.2. Vocabulary

The vocabulary of a language model is the set of unique subword units, called tokens, recognized by the model. Models such as PEGASUS and mT5 construct the vocabulary by training a subword tokenizer that aims to learn a favorable separation of an input text. As in PEGASUS, we train the language-independent SentencePiece unigram tokenizer (Kudo and Richardson, 2018) to construct a vocabulary for Gullfaxi. We use the same configuration as for PEGASUS, a vocabulary size of 96k, and do not differentiate between lowercase and uppercase letters. The Gullfaxi tokenizer is trained on documents from the IGC. The mT5 model comes with a multilingual vocabulary of size 250k, obtained from training a SentencePiece tokenizer on the mC4 dataset.

### 3. Methods

#### 3.3.3. Hyperparameter configuration

Gullfaxi is computationally expensive to train. Therefore, we do not conduct a search for optimal hyperparameters. Unless stated explicitly, we use the same configuration as PEGASUS, which we trust to be a good first choice. Further details of our experiments’ hyperparameters and training configuration can be found in Appendix B.

The training was carried out on a high-performance computing cluster at the Jülich Supercomputing Centre using multiple GPUs and a distributed configuration implemented with the HuggingFace Accelerate library. All models are implemented with the HuggingFace Transformers library. Gullfaxi is trained from scratch and its implementation uses the same configuration as PEGASUS. The configuration of the mT5 model is loaded from the `google/mt5-base` repository, using pre-trained weights and the corresponding tokenizer.

Gullfaxi<sub>LARGE</sub> is pre-trained with a step batch size of 8192 examples and Gullfaxi<sub>BASE</sub> is pre-trained with a batch size of 256. For fine-tuning all models, the batch size is set to 256 examples. The maximum input and output tokens are respectively 512 and 256 for pre-training but vary for fine-tuning tasks. For optimization, we use AdaFactor (Shazeer and Stern, 2018) with an internally adjusted learning rate and cross-entropy loss. During fine-tuning, we use a label-smoothed regularization with a value of 0.1. At test time, we use beam size 8 with a length penalty of 0.8 for all tasks.

#### 3.3.4. Downstream tasks

We perform evaluations on a few downstream summarization tasks. Particularly, we create three fine-tuning tasks from the RRN dataset testing different abilities for abstractive summarization. The RRN dataset is randomly split into a train (60%), a validation (20%), and a test set (20%) for each fine-tuning task. The RRN downstream tasks are listed below.

Intro + Main → Summary Producing the summary from the intro along with the main part of the story. Although the intro and the summary can be completely different, the intro is often similar or the same as the summary. Therefore, this fine-tuning task is somewhat related to extractive summarization and we expect that the model might learn to copy sentences from the intro.

Main → Intro Generating the intro from the main part. The intro often outlines the subject of the main part in one or two sentences. The intro and the main text rarely share the same sentences and for that reason we expect the model to become more abstractive than in the previous scenario. However, note that information in the intro may not appear in the main part. Due to the similarity of the intro and the summary, we expect this task to be similar to the Main → Summary task but since there are more intros than summaries in the dataset, we prefer this task.

Intro → Title Producing the title of the story from the intro. The title is much shorter compared to the target text in previous tasks and only includes the keywords from the story. We expect this to also have a more abstractive output but if titles have some repetitive patterns the model might overfit them.

The preceding RRN tasks are based on a limited number of data examples for training. To evaluate the performance of our approach on larger corpora, we fine-tune Gullfaxi on the Icelandic translations of the XSum and CNN/DailyMail datasets and compare the results to those of the PEGASUS study. Additionally, we continue fine-tuning the models on the RRN tasks after fine-tuning them on the translations, as we believe that this data augmentation approach may improve the generalization of the summarization task by providing more diversity and training examples. However, it’s worth noting that machine translations can sometimes be unnatural or grammatically incorrect, and this may affect the quality of the summaries.

## 3.4. Performance

### 3.4.1. ROUGE

The Recall-Oriented Understudy for Gisting Evaluation (ROUGE) scoring algorithm is the standard evaluation metric for automatic summarization. We use it to calculate the similarity between the model’s summary and a reference summary.

We use  $\text{count}_{\text{match}}(\text{gram}_n)$  to represent the number of matching n-grams and similarly,  $\text{count}_{\text{ref}}(\text{gram}_n)$  and  $\text{count}_{\text{model}}(\text{gram}_n)$  refer to the number of n-grams in the reference and the model output relatively. Referring to ROUGE-N as the number of overlapping n-grams between the two texts, we can obtain the precision, recall, and F1-score as follows:

$$\text{ROUGE-N precision} = \frac{\text{count}_{\text{match}}(\text{gram}_n)}{\text{count}_{\text{model}}(\text{gram}_n)}$$

### 3. Methods

$$\text{ROUGE-N recall} = \frac{\text{count}_{\text{match}}(\text{gram}_n)}{\text{count}_{\text{ref}}(\text{gram}_n)}$$

$$\text{F1-score} = 2 \times \frac{\text{precision} \times \text{recall}}{\text{precision} + \text{recall}}$$

Furthermore, ROUGE-L uses the longest common subsequence (LCS) to score the similarity between the documents.  $\text{LCS}(\text{gram}_1)$  represents the longest sequence of words that are shared between the reference and the model output. Note that the sequence does not have to be consecutive. We define the ROUGE-L metrics as follows:

$$\text{ROUGE-L precision} = \frac{\text{LCS}(\text{gram}_1)}{\text{count}_{\text{model}}(\text{gram}_1)}$$

$$\text{ROUGE-L recall} = \frac{\text{LCS}(\text{gram}_1)}{\text{count}_{\text{ref}}(\text{gram}_1)}$$

For evaluating the results in this study we use the F1-score of ROUGE-1, ROUGE-2, and ROUGE-L. These are the set of metrics that are the most commonly used for summarization tasks and can be used for comparison to other studies.

#### 3.4.2. Human Evaluation

For a better understanding of the model performance, we randomly sample a subset of summary outputs for human evaluation. The purpose is to evaluate the quality of the summaries and determine their correctness as this is not specifically accounted for with the ROUGE scores. We let a single annotator rate the summaries on three criteria, where 1 means that a criterion is met but 0 that it is not.

- **Relevance.** The output is relevant to the reference text and relates to the subject of the story.
- **Correctness.** The output is consistent with the reference text, factually accurate, and does not add unrelated information.
- **Language.** The output is grammatically correct and natural, not repeating phrases or using non-Icelandic words.

# 4. Results

## 4.1. Model Comparison

We evaluate the performance of the two Gullfaxi models and mT5 on the RRN dataset. We present the results in Table 4.1, which lists the average ROUGE F1-scores (R1/R2/RL) of the fine-tuned models for a corresponding task. Our results show that mT5<sub>BASE</sub> consistently outperforms the Gullfaxi models on all tasks and Gullfaxi<sub>BASE</sub> performs better than Gullfaxi<sub>LARGE</sub>. The difference between models is particularly prominent where we might expect that an extractive behavior achieves a high score as in the first task (Intro + Main → Summary). We provide examples of the model outputs and their scores in Appendix C.

Table 4.1: A comparison of Gullfaxi<sub>LARGE</sub>, Gullfaxi<sub>BASE</sub>, and mT5<sub>BASE</sub> fine-tuned on the RRN dataset. The scores listed are the average ROUGE F1-scores (R1/R2/RL).

Model	Intro + Main → Summary	Main → Intro	Intro → Title
Gullfaxi <sub>BASE</sub>	29.44/10.18/26.81	20.51/5.70/18.45	26.62/5.82/26.59
Gullfaxi <sub>LARGE</sub>	23.00/6.02/21.01	18.58/4.79/16.81	25.40/5.09/25.37
mT5 <sub>BASE</sub>	54.86/38.78/52.13	24.80/11.21/22.99	27.07/5.12/26.81

By manually reviewing the text outputs of the models we find that mT5 almost completely relies on an extractive approach in the first task (Intro + Main → Summary), simply copying the intro, resulting in a much higher score compared to Gullfaxi. In the other tasks, we see somewhat abstractive output from all models. Other attributes that we notice that result in lower ROUGE scores include repetition, different declensions than in the reference summary, and different lengths of the output.

Comparing the fine-tuning time we found Gullfaxi adapting faster to the summarization task, achieving a considerably higher score at the start compared to mT5. For pre-training, Gullfaxi<sub>LARGE</sub> required considerably more computational resources per step due to larger batch size and more trainable parameters compared

## 4. Results

to Gullfaxi<sub>BASE</sub>. Figure 4.1 shows the pre-training validation loss for both models. Gullfaxi<sub>LARGE</sub> has a faster rate of convergence but note that it traverses more examples each training step. Due to limited computing resources available for the project, the larger model was not trained to convergence but still slightly surpassed the best base model checkpoint.

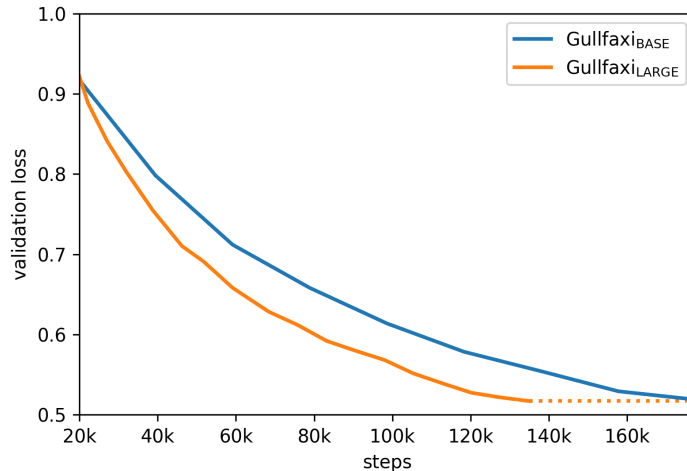


Figure 4.1: The pre-training validation loss model comparison. Gullfaxi<sub>LARGE</sub> and Gullfaxi<sub>BASE</sub> were pre-trained on the IGC using batch sizes 8192 and 256 respectively.

## 4.2. Low-resource Fine-tuning

To further study the low-resource summarization setting we compared the fine-tuning of Gullfaxi<sub>BASE</sub> and mT5<sub>BASE</sub> using the first  $10^k$  ( $k = 1, 2, 3$ ) examples from the RRN training set. During the low-resource fine-tuning, we used all the validation set examples and at test time we used all the test set examples. Figure 4.2 shows the results of the low-resource fine-tuning. The dashed lines represent the performance of Gullfaxi<sub>BASE</sub> and the solid lines represent the performance of mT5<sub>BASE</sub>. The transparent lines represent the performance of fine-tuning a randomly initialized model with the same architecture as Gullfaxi<sub>BASE</sub> on the full RRN training set without any pre-training, we refer to this model as Transformer<sub>BASE</sub>.

We can see from Figure 4.2 that even without any fine-tuning, Gullfaxi<sub>BASE</sub> performs better than Transformer<sub>BASE</sub> on some tasks. Note that the mT5<sub>BASE</sub> model used has only been pre-trained on mC4 excluding any supervised training and therefore using it without any fine-tuning naturally results in a poor summarization performance. When the number of training examples is increased, mT5 gradually exceeds Gullfaxi

in all tasks. Using all training examples we see both models achieving significantly higher scores than  $\text{Transformer}_{\text{BASE}}$ .

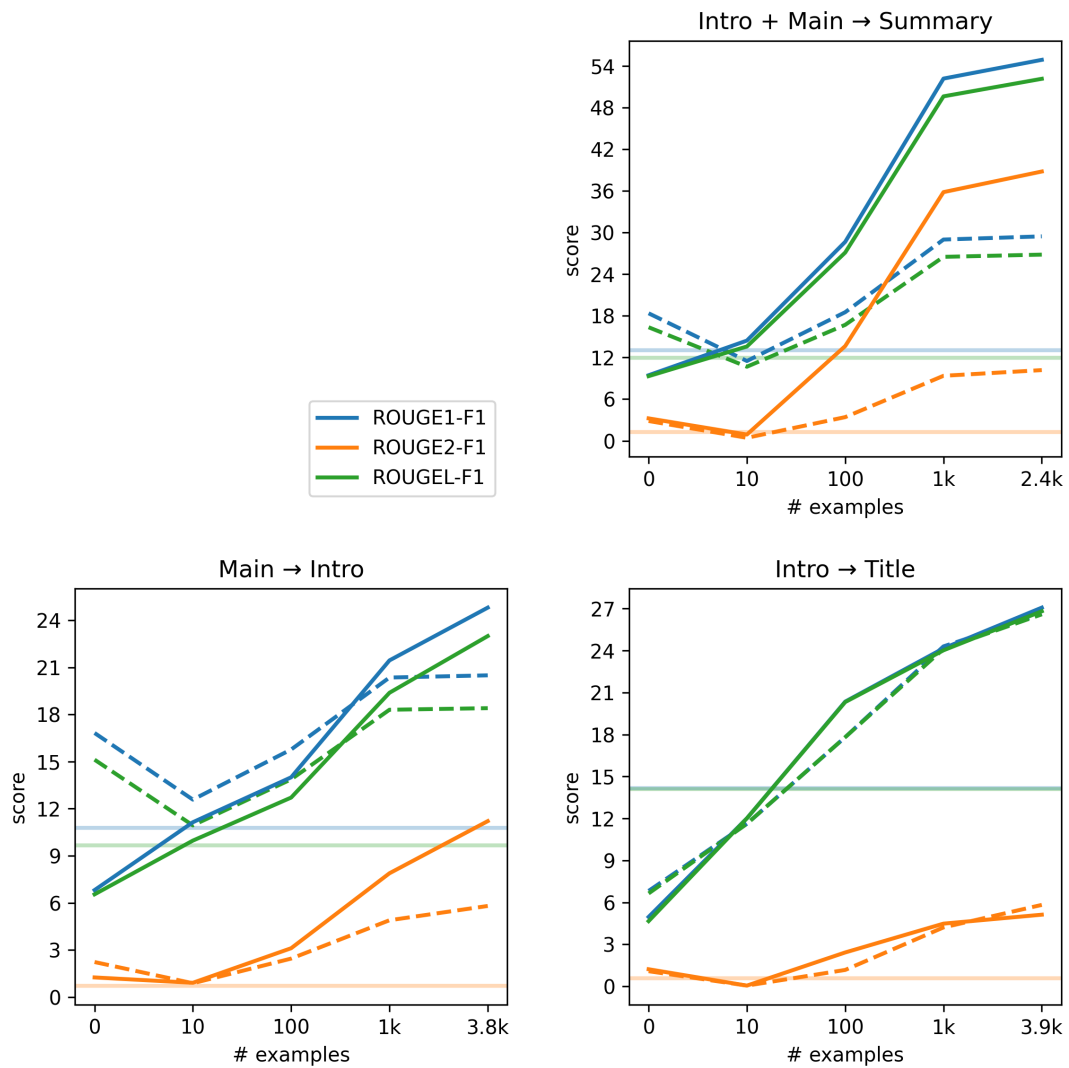


Figure 4.2: The performance of  $mT5_{\text{BASE}}$  (solid lines) and  $Gullfaxi_{\text{BASE}}$  (dashed lines) on the RRN tasks with a limited number of fine-tuning examples. The transparent horizontal lines represent the performance of the same capacity  $Gullfaxi_{\text{BASE}}$  model,  $\text{Transformer}_{\text{BASE}}$ , trained on the task with all training examples but without any pre-training.



### 4.3. Fine-tuning Data Augmentation

To investigate the potential benefits of data augmentation, we fine-tuned Gullfaxi<sub>BASE</sub> on the Icelandic translations of XSum and CNN/DailyMail. For the two lengthy output RRN tasks, we evaluated the model’s performance when the translated datasets were included in the fine-tuning process, either alone or followed by fine-tuning on the corresponding RRN tasks. The ROUGE scores with and without fine-tuning are presented in Table 4.2. By comparing the results to those in Table 4.1, we can see that there is a limited generalization to the summarization task when the RRN data is not included in the training. However, when the translations are combined with RRN, the scores are higher, particularly when using the CNN/DailyMail dataset.

Table 4.2: Gullfaxi<sub>BASE</sub>’s performance on RRN when the Icelandic translations of XSum and CNN/DailyMail are used for fine-tuning, either exclusively or followed by fine-tuning on the corresponding RRN task. The ROUGE F1-scores (R1/R2/RL) are listed in the table.

Model	Intro + Main → Summary	Main → Intro
Gullfaxi <sub>BASE</sub> (no fine-tuning)	18.36/2.85/16.35	16.82/2.24/15.12
Gullfaxi <sub>BASE</sub> (XSum)	17.26/2.96/15.17	15.82/2.18/13.73
Gullfaxi <sub>BASE</sub> (CNN/DailyMail)	26.49/8.98/24.21	17.76/3.54/16.02
Gullfaxi <sub>BASE</sub> (XSum + RRN)	29.40/9.29/26.41	20.66/5.19/18.54
Gullfaxi <sub>BASE</sub> (CNN/DailyMail + RRN)	42.50/21.32/39.62	22.20/6.24/19.73

The results of the ROUGE evaluation indicate that using the XSum dataset alone does not improve performance on the RRN task. However, it should be noted that the XSum target texts used in training are shorter than those used in RRN tasks. Our fine-tuning results show improvement over no fine-tuning when Gullfaxi is fine-tuned on CNN/DailyMail, indicating that there may be some generalization of the summarization task. Additionally, when fine-tuned in combination with RRN, the scores are significantly higher than without the CNN/DailyMail translations. Upon manual inspection of the model’s outputs, we observed an increase in grammatical errors when using the translations for fine-tuning.

In order to highlight the contrast in ROUGE scores between Icelandic and English, we evaluated the performance of our models fine-tuned on XSum and CNN/DailyMail on their respective test sets. Table 4.3 presents the results for Gullfaxi, both with and without fine-tuning, as well as a comparison to the PEGASUS model’s English data scores from the original study. It is apparent that fine-tuning leads to a notable improvement in performance on both datasets. However, when comparing Gullfaxi to PEGASUS, it is evident that the PEGASUS model’s scores for English are significantly higher.

Table 4.3: Gullfaxi<sub>BASE</sub>’s ROUGE F1-scores (R1/R2/RL) with and without fine-tuning on the Icelandic translations of XSum and CNN/DailyMail. The scores listed for PEGASUS<sub>BASE</sub> and PEGASUS<sub>LARGE</sub> are the highest English language fine-tuning scores obtained from the PEGASUS paper.

Model	XSum <sub>is</sub>	CNN/DailyMail <sub>is</sub>
Gullfaxi <sub>BASE</sub> (no fine-tuning)	13.31/1.12/11.41	13.13/1.27/12.11
Gullfaxi <sub>LARGE</sub> (no fine-tuning)	13.19/1.19/11.42	13.34/1.25/12.45
Gullfaxi <sub>BASE</sub> (XSum)	23.46/7.31/19.86	-
Gullfaxi <sub>LARGE</sub> (XSum)	22.75/6.88/19.34	-
Gullfaxi <sub>BASE</sub> (CNN/DailyMail)	-	24.61/7.74/23.14
Gullfaxi <sub>LARGE</sub> (CNN/DailyMail)	-	23.81/6.65/22.32
	XSum <sub>en</sub>	CNN/DailyMail <sub>en</sub>
PEGASUS <sub>BASE</sub>	39.79/16.58/31.70	41.79/18.81/38.93
PEGASUS <sub>LARGE</sub>	47.21/24.56/39.25	44.17/21.47/41.11

## 4.4. Human Evaluation

To further evaluate performance, we manually annotated the output summaries of the models based on the criteria described in Section 3.4.2. For evaluation, we sampled at random 50 examples of the Main  $\rightarrow$  Intro task, testing the ability to generate a few-sentence abstractive summary. In Table 4.4 we compare the average scores for Gullfaxi<sub>BASE</sub> and mT5. Examples of the human annotations along with the model output can be found in Appendix D.

Table 4.4: The human evaluation scores for the Main  $\rightarrow$  Intro task. The scores listed are the average of 50 annotations for each model.

Model	Relevance	Correctness	Language
Gullfaxi <sub>BASE</sub> (RRN)	0.74	0.08	0.42
Gullfaxi <sub>BASE</sub> (CNN/DailyMail)	0.46	0.08	0.06
Gullfaxi <sub>BASE</sub> (CNN/DailyMail + RRN)	0.64	0.04	0.10
mT5 <sub>BASE</sub> (RRN)	0.84	0.46	0.54

Overall, our evaluation revealed that the output intros generated by all models were of lower quality than the intros written by humans. The model outputs were often pertinent to the reference text but failed to effectively summarize it. The mT5 model exhibited the best performance across all criteria. However, we observed that mT5 tended to extract bits from the reference text rather than generating novel phrases, which could partially explain its higher relevance and correctness scores compared to Gullfaxi. Fine-tuning Gullfaxi on the Icelandic translations of CNN/DailyMail

#### 4. Results

resulted in the worst performance, particularly with regard to its ability to correctly use Icelandic grammar. This is evident from the results presented in Table 4.5, which shows examples of incorrect inflections of words.

Table 4.5: Examples of ungrammatical output text of *Gullfaxi*<sub>BASE</sub> fine-tuned exclusively on the Icelandic translations of CNN/DailyMail. Corrections are in parentheses. The inflections of the words in red are incorrect.

---

<b>Reference summary</b>	Frankvæmdastjóri Vistorku á Akureyri segir raunsæjan kost að Ísland geti orðið algjörlega óháð olíu á næstu árum og þar með sjálfbært um alla orkuframleiðslu. Heildræna stefnu vanti þó í málaflokknum.
<b>Model output</b>	frankvæmdastjóri segir að það <b>er engin</b> (séu engar) hindranir <b>til</b> (fyrir því) að flýta banni við innflutningi á olíu
<b>Reference summary</b>	Í fyrsta sinn í 15 ár er stefnt að því að byggja fjölda íbúðarhúsa norður af Akureyri. Mikill áhugi er á lóðunum og færri fengu úthlutun en vildu.
<b>Model output</b>	á síðustu 19 árum <b>hafa lóðir</b> (hefur lóðum) verið úthlutað til eldri borgara. áhugi á <b>lóðir</b> (lóðum) hefur aukist á undanförunum árum.

---

## 5. Discussion

In this study, we explored techniques for taking on a challenging low-resource NLP task, creating an abstractive summarization model for Icelandic. We tested some well-known approaches for training such models, particularly regarding a data-scarce setting. We discovered the limitations of our work as well as possible improvements for future work. In the following sections, we discuss the results of our research, implications, and future work.

### 5.1. Abstractive Summarization Data

Before this study, there did not exist an abstractive summarization corpus for Icelandic. Manually creating such datasets is labor intensive and that might explain the lack of low-resource studies on this subject compared to many other NLP tasks. However, summary-like content can be found widely on the internet, presenting opportunities for the automatic collection of such data for the training of summarization models within particular domains.

For the project, we collected a news domain abstractive summarization dataset for Icelandic, RÚV Radio News (RRN), and demonstrated that it can be used for training and evaluating summarization models. We created three different tasks from the stories in RRN that test different summarization learning skills and vary in terms of length and abstractiveness.

While the RRN story segments were not specifically written as summaries, they proved suitable for our experiment as they closely resemble summaries in nature. However, for that reason and the limited diversity of the data, we expect that a model trained on RRN might not generalize well to a different summarization setting and/or domain. In particular, examples from the news domain have a different structure compared to other domains as the key information usually appears at the beginning which naturally causes the model to focus specifically on that part. That could lead to overfitting when applying the model to other datasets where the critical information does not necessarily appear at the beginning of the text.

## 5. Discussion

The examples included in RRN are only from the most recent years and the dataset is small compared to the commonly used summarization datasets such as XSum and CNN/DailyMail. It should be noted that the examples in RRN exist many years back and further collection and processing could be conducted. However, the collection and cleaning of the texts in RRN were rather time-consuming as the consistency in the format of the transcripts was lacking. Further development of language resources for Icelandic should also include making publicly funded organizations such as RÚV more aware of their data processing as their work could have a direct effect on the development of NLP in Iceland.

Further advances in summarization solutions for Icelandic will likely require more summarization data. RRN only covers the news domain but practical abstractive summarization solutions will require greater training data diversity. However, RRN is a starting point for further development of automatic abstractive summarization solutions for Icelandic and we encourage researchers in the field to explore its usage.

### 5.2. Pre-training of Gullfaxi

For the pre-training of Gullfaxi, we applied a self-supervised pre-training objective that resembles the summarization task in the hope of creating a model that can adapt quickly when fine-tuned for summarization as was the case for English. From our results, we found that there were considerable benefits of the pre-training when comparing Gullfaxi to a model that had not been pre-trained. We found Gullfaxi adapting quickly to the RRN tasks compared to mT5 and inclining towards a more abstractive behavior when fine-tuned. Our results demonstrate the importance of pre-training in a low-resource setting for achieving syntactic awareness and context understanding before fine-tuning.

Comparing the pre-training of Gullfaxi to PEGASUS there are a few differences to consider. For pre-training, Gullfaxi uses the IGC which consists of circa 5 million documents whereas PEGASUS was pre-trained on C4 (350M web pages) and HugeNews (1.5B news-like articles). The difference in the volume of data is apparent and we expect this to have an impact on the results of our study. Studies have shown that large language models tend to benefit from large volumes of pre-training data (Devlin et al., 2019; Hoffmann et al., 2022; Raffel et al., 2019) and we expect that our models would do so too. In general, further development of large Icelandic language corpora is important for continued advancements.

But with limited data, does parameter count matter? We compared the pre-training of the two Gullfaxi models and found that the larger model had lower ROUGE scores in all RRN tasks. Gullfaxi<sub>LARGE</sub> had more than twice the number of trainable pa-

rameters of Gullfaxi<sub>BASE</sub> and required considerably more training time. The larger model used a much larger batch size but due to the limited size of the IGC, both models cover the same amount of data examples. This is different from the PEGASUS study where the larger model traverses more data. However, note that Gullfaxi<sub>LARGE</sub> was not trained to convergence but our results suggest that there is little benefit in more training as the slope of the validation loss is almost flat at the end of the training period. Still, that does not mean that the model cannot be improved with further training but the return on computational time in terms of reduced loss is diminished.

Our results suggest that there was no benefit in using the larger model compared to the base model on a fixed computational budget. The reason could be that the benefit of the larger model is constrained by the number of pre-training examples or that further training time would be necessary for it to improve. Keep in mind that the optimal hyperparameters can also be different between English and Icelandic and since we based our hyperparameter selection on the PEGASUS study, it could influence our results. The optimal hyperparameter choice can further depend on the amount of training data as other studies have demonstrated (Hoffmann et al., 2022). Note that we chose the same parameters due to the cost of evaluating them. It is thus an open question whether hyperparameter choices differ between languages where the amount of training data is the same.

## 5.3. Low-resource Languages Benefit from Multilingual Pre-training

In this study, we fine-tuned both Gullfaxi and the multilingual mT5 model on the RRN dataset and compared their performance. Both Gullfaxi<sub>BASE</sub> and mT5<sub>BASE</sub> share a similar Transformer model architecture, but the primary difference is in their pre-training phase. Our experiments revealed that mT5 performed better overall than Gullfaxi, achieving a higher ROUGE score in all tasks and receiving a higher rating in the human evaluation. However, we also noticed that mT5 often extracted phrases rather than generating new ones, which is not ideal for abstractive summarization.

There is a considerable difference between the pre-training data of mT5 compared to the IGC. The mT5 model is pre-trained on 6.6B pages in various languages, including 2.1M Icelandic pages. Although the Icelandic data is less than for Gullfaxi, we expect the general knowledge acquired with multilingual pre-training to translate between languages when fine-tuned for tasks such as summarization. A few Icelandic studies (Snæbjarnarson and Einarsson, 2022a; Snæbjarnarson et al., 2022) have

reached the same conclusion regarding the benefits of using multilingual models and for that reason, it is worthwhile to consider the possible advantages of using such models for low-resource languages.

### 5.4. The Value of Adding Fine-tuning Examples

We evaluated the performance of our models in an extremely low-resource summarization setting by limiting the amount of fine-tuning data. We found that the zero-shot performance of Gullfaxi<sub>BASE</sub> outperformed Transformer<sub>BASE</sub> in two out of the three RRN tasks. Additionally, when using the full supervised datasets, Gullfaxi and mT5 achieved significantly higher scores than Transformer<sub>BASE</sub>. This performance demonstrates the advantages of pre-training when training for low-resource summarization tasks. Examining the learning curves in Figure 4.2, we can infer that the models would benefit from more fine-tuning data, but the returns would decrease. The models benefit greatly from only 1k examples, emphasizing the importance of creating small summarization datasets within specific domains.

### 5.5. Limitations of the Approach

All the models tended to generate summaries some of the time that were inconsistent with the source text. Although they were often relevant to the subject, we found them to add incorrect statements. This is a common problem for abstractive summarization models (Cao et al., 2017) as well as for other generative models (Bender et al., 2021) and limits the practical usage of such models. In that context, we wonder if the pre-training objective might encourage such behavior as the model is trained to produce information that does not necessarily appear in the original text. Constraining the model to behave properly and generate relevant output is an ongoing challenge. Currently, the most successful approach seems to be reinforcement learning with human feedback that OpenAI has applied to train their Instruct GPT model (Ouyang et al., 2022).

State-of-the-art models and breakthrough studies in NLP commonly use high-resource languages such as English. One can wonder to what extent the choice of language impacts the result when the training budget and amount of training data are fixed. This brings up a question regarding the general performance that can be achieved across different languages for state-of-the-art models. From our results, we cannot conclude that Icelandic is more challenging than English but training models on the same amount of English data could reveal whether the difference is large or small.

## 5.6. Machine Translations for Data Augmentation

For fine-tuning data augmentation, we used a state-of-the-art machine translation model to increase the amount of summarization data by translating two relatively large English language summarization datasets to Icelandic. Although we knew that the translations might not be perfect, we hoped that the grammar knowledge acquired from the pre-training would repress inaccurate translations in the fine-tuning of the model. We expected that the model would benefit from a large fine-tuning dataset for learning the generalization of the summarization task.

Our results revealed that fine-tuning Gullfaxi on the Icelandic translations led to models that generated a significant number of grammatical errors. In the human evaluation, the models performed notably poorly in terms of language and relevance, and the generated outputs often had unnatural phrasing. This outcome highlights the critical role that the quality of the training data plays in NLP and underscores the need for further data collection efforts for Icelandic text.

When reviewing the Icelandic translations there are a few things to note. Although most of them are easily comprehensible for a native speaker, they tend to be unnatural, use unusual wording, and often have the wrong inflection of words. For that reason, we think that the translations might not be as easily comprehended by the model since the use of the language is quite different from the pre-training data. That does not mean that this augmentation is not helpful since it could be used in some other way to improve the model’s performance. For example, it could be mixed with RRN such that RRN examples have a higher weight in the loss function than the translated examples. An experiment might reveal that a non-zero weight for translated examples yields a performance superior to zero weights (i.e. leaving them out).

## 5.7. Evaluating Abstractive Summaries

We evaluated the quality of the output summaries generated by the models using ROUGE scores as well as scores from a human evaluation. Evaluating abstractive summaries automatically is challenging, as high-quality abstractive summaries can take many different forms, including the use of novel words and paraphrasing sentences from the input text. Manually evaluating such summaries is laborious, as it requires a thorough understanding of the reference text and the ability to identify key information, followed by an evaluation of the model’s summary based on established criteria. To establish a reliable baseline, it would be beneficial to compute these metrics on the RRN dataset.



## 5. Discussion

The ROUGE scoring algorithm is a heuristic for evaluating the quality of summarization. For example, a good abstractive summary can have a low ROUGE score if its distribution of  $n$ -grams is not the same as in the reference summary. ROUGE is known to favor lexical similarity which can make it unsuitable for evaluating abstractive summaries (Ng and Abrecht, 2015). This is especially noteworthy when dealing with more complex languages such as Icelandic. Icelandic nouns can take sixteen different forms and verbs and adjectives can take more than a hundred forms. Therefore words with the same meaning can be seen as completely different by ROUGE.

Comparing the ROUGE scores of model summaries in this study to English language studies, we observe that they are generally much lower. We hypothesize that this can somewhat be explained by the differences in the languages' grammar. For that reason, it is hard to compare the scores in this study to other studies and makes us wonder if the ROUGE scores are generally a good evaluation metric for abstractive summarization. However, it should be mentioned that we expect that calculating the ROUGE scores from the lemmatized Icelandic sentences could somewhat reduce the differences in scoring between languages and should be considered for further evaluation of abstractive summarization in Icelandic.

For this study, we found the ROUGE scores not representing the quality of the summaries very well as they did not account for whether the texts were extractive or abstractive, had wrong grammar, or were consistent with the reference text. In this regard, we performed a human evaluation of the models' summaries to better understand their quality and particularly study the different characteristics of the models.

The human evaluation revealed a number of things that were not inferable from the ROUGE score. As discussed in the preceding sections, the summaries tended to be factually inaccurate and the quality of the grammar varied by models. The summaries that met all the evaluation criteria were usually high-quality although it should be noted that a summary can meet all criteria without properly summarizing the input text. Ranking or scoring summaries where a human-written summary is among them could be a good way to evaluate the performance further in relation to human-written summaries.

## 5.8. Conclusion

The object of our work was to study methods for automatic abstractive summarization in a low-resource setting. In particular, we applied current state-of-the-art models to Icelandic text, which is morphologically rich and does not possess the

same amount of data as has been used to train these models for languages such as English. We collected and published the first publicly available abstractive summarization dataset for Icelandic and used it for training and evaluation of our models.

Based on our findings there are significant benefits to using multilingual pre-training in this setting. When fine-tuned for summarization, the multilingual mT5 consistently outperformed a similar capacity PEGASUS model pre-trained from scratch on Icelandic text only. We found that using machine translations to increase the amount of fine-tuning data resulted in lower-quality summaries that were ungrammatical and more inconsistent with the reference text.

For future work, we see opportunities for further collection of abstractive summarization data for Icelandic. The RRN dataset can be expanded and we believe that summary-like content can be found in various domains such as scholarly works, legal texts, and medical documents. We emphasize the benefits of using the pre-trained mT5 model and we expect it to apply to other generative tasks and different languages. Automatic evaluation of abstractive summarization models is an open problem and further study of metrics that might be better suited for Icelandic text should be conducted.



# Bibliography

- Rami Al-Rfou, Dokook Choe, Noah Constant, Mandy Guo, and Llion Jones. Character-Level Language Modeling with Deeper Self-Attention. *Proceedings of the AAAI Conference on Artificial Intelligence*, 33(01):3159–3166, July 2019. ISSN 2374-3468. doi: 10.1609/aaai.v33i01.33013159. URL <https://ojs.aaai.org/index.php/AAAI/article/view/4182>. Number: 01. 4
- Dzmitry Bahdanau, Kyunghyun Cho, and Yoshua Bengio. Neural machine translation by jointly learning to align and translate. In Yoshua Bengio and Yann LeCun, editors, *3rd International Conference on Learning Representations, ICLR 2015, San Diego, CA, USA, May 7-9, 2015, Conference Track Proceedings, 2015*. URL <http://arxiv.org/abs/1409.0473>. 3
- Iz Beltagy, Matthew E. Peters, and Arman Cohan. Longformer: The Long-Document Transformer. *arXiv:2004.05150 [cs]*, December 2020. URL <http://arxiv.org/abs/2004.05150>. arXiv: 2004.05150. 4
- Emily M. Bender, Timnit Gebru, Angelina McMillan-Major, and Shmargaret Shmitchell. On the Dangers of Stochastic Parrots: Can Language Models Be Too Big? In *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency, FAccT '21*, pages 610–623, New York, NY, USA, March 2021. Association for Computing Machinery. ISBN 978-1-4503-8309-7. doi: 10.1145/3442188.3445922. URL <https://doi.org/10.1145/3442188.3445922>. 24
- Tom B. Brown, Benjamin Mann, Nick Ryder, Melanie Subbiah, Jared Kaplan, Prafulla Dhariwal, Arvind Neelakantan, Pranav Shyam, Girish Sastry, Amanda Askell, Sandhini Agarwal, Ariel Herbert-Voss, Gretchen Krueger, Tom Henighan, Rewon Child, Aditya Ramesh, Daniel M. Ziegler, Jeffrey Wu, Clemens Winter, Christopher Hesse, Mark Chen, Eric Sigler, Mateusz Litwin, Scott Gray, Benjamin Chess, Jack Clark, Christopher Berner, Sam McCandlish, Alec Radford, Ilya Sutskever, and Dario Amodei. Language Models are Few-Shot Learners, July 2020. URL <http://arxiv.org/abs/2005.14165>. arXiv:2005.14165 [cs]. 4
- Ziqiang Cao, Furu Wei, Wenjie Li, and Sujian Li. Faithful to the Original: Fact Aware Neural Abstractive Summarization, November 2017. URL <http://arxiv.org/abs/1711.04434>. arXiv:1711.04434 [cs]. 24

## BIBLIOGRAPHY

- Sumit Chopra, Michael Auli, and Alexander M. Rush. Abstractive Sentence Summarization with Attentive Recurrent Neural Networks. In *Proceedings of the 2016 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, pages 93–98, San Diego, California, June 2016. Association for Computational Linguistics. doi: 10.18653/v1/N16-1012. URL <https://aclanthology.org/N16-1012>. 3
- Karin Christiansen. *Summarization of Icelandic Texts*. PhD thesis, Master’s thesis, Reykjavik University, Reykjavik, Iceland, 2014. 6
- Alexis Conneau, Kartikay Khandelwal, Naman Goyal, Vishrav Chaudhary, Guillaume Wenzek, Francisco Guzmán, Edouard Grave, Myle Ott, Luke Zettlemoyer, and Veselin Stoyanov. Unsupervised Cross-lingual Representation Learning at Scale, April 2020. URL <http://arxiv.org/abs/1911.02116>. arXiv:1911.02116 [cs]. 6
- Jón Daðason, Hrafn Loftsson, Salome Sigurðardóttir, and Þorsteinn Björnsson. IceSum: An Icelandic Text Summarization Corpus. In *Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Student Research Workshop*, pages 9–14, Online, June 2021. Association for Computational Linguistics. doi: 10.18653/v1/2021.naacl-srw.2. URL <https://aclanthology.org/2021.naacl-srw.2>. 6
- Jón Friðrik Daðason and Hrafn Loftsson. Pre-training and Evaluating Transformer-based Language Models for Icelandic. In *Proceedings of the Thirteenth Language Resources and Evaluation Conference*, pages 7386–7391, Marseille, France, June 2022. European Language Resources Association. URL <https://aclanthology.org/2022.lrec-1.804>. 6
- Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. *arXiv:1810.04805 [cs]*, May 2019. URL <http://arxiv.org/abs/1810.04805>. arXiv: 1810.04805. 4, 5, 22
- Li Dong, Nan Yang, Wenhui Wang, Furu Wei, Xiaodong Liu, Yu Wang, Jianfeng Gao, Ming Zhou, and Hsiao-Wuen Hon. Unified Language Model Pre-training for Natural Language Understanding and Generation. *arXiv:1905.03197 [cs]*, October 2019. URL <http://arxiv.org/abs/1905.03197>. arXiv: 1905.03197. 5
- Marzieh Fadaee, Arianna Bisazza, and Christof Monz. Data Augmentation for Low-Resource Neural Machine Translation. In *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers)*, pages 567–573, 2017. doi: 10.18653/v1/P17-2090. URL <http://arxiv.org/abs/1705.00440>. arXiv:1705.00440 [cs]. 6
- Karl Moritz Hermann, Tomas Kocisky, Edward Grefenstette, Lasse Espeholt, Will Kay, Mustafa Suleyman, and Phil Blunsom. Teaching Machines to Read and

- Comprehend. In *Advances in Neural Information Processing Systems*, volume 28. Curran Associates, Inc., 2015. URL <https://proceedings.neurips.cc/paper/2015/hash/afdec7005cc9f14302cd0474fd0f3c96-Abstract.html>. 8
- Jordan Hoffmann, Sebastian Borgeaud, Arthur Mensch, Elena Buchatskaya, Trevor Cai, Eliza Rutherford, Diego de Las Casas, Lisa Anne Hendricks, Johannes Welbl, Aidan Clark, Tom Hennigan, Eric Noland, Katie Millican, George van den Driessche, Bogdan Damoc, Aurelia Guy, Simon Osindero, Karen Simonyan, Erich Elsen, Jack W. Rae, Oriol Vinyals, and Laurent Sifre. Training Compute-Optimal Large Language Models, March 2022. URL <http://arxiv.org/abs/2203.15556>. arXiv:2203.15556 [cs]. 2, 22, 23
- Taku Kudo and John Richardson. SentencePiece: A simple and language independent subword tokenizer and detokenizer for Neural Text Processing. *arXiv:1808.06226 [cs]*, August 2018. URL <http://arxiv.org/abs/1808.06226>. arXiv: 1808.06226. 11
- Mike Lewis, Yinhan Liu, Naman Goyal, Marjan Ghazvininejad, Abdelrahman Mohamed, Omer Levy, Ves Stoyanov, and Luke Zettlemoyer. BART: Denoising Sequence-to-Sequence Pre-training for Natural Language Generation, Translation, and Comprehension. *arXiv:1910.13461 [cs, stat]*, October 2019. URL <http://arxiv.org/abs/1910.13461>. arXiv: 1910.13461. 5
- Peter J. Liu, Mohammad Saleh, Etienne Pot, Ben Goodrich, Ryan Sepassi, Lukasz Kaiser, and Noam Shazeer. Generating Wikipedia by Summarizing Long Sequences. *arXiv:1801.10198 [cs]*, January 2018. URL <http://arxiv.org/abs/1801.10198>. arXiv: 1801.10198. 4
- Alexandre Magueresse, Vincent Carles, and Evan Heetderks. Low-resource Languages: A Review of Past Work and Future Challenges, June 2020. URL <http://arxiv.org/abs/2006.07264>. arXiv:2006.07264 [cs]. 6
- Tomas Mikolov, Kai Chen, Greg Corrado, and Jeffrey Dean. Efficient Estimation of Word Representations in Vector Space, September 2013. URL <http://arxiv.org/abs/1301.3781>. arXiv:1301.3781 [cs]. 5
- Ramesh Nallapati, Bowen Zhou, Cicero Nogueira dos santos, Caglar Gulcehre, and Bing Xiang. Abstractive Text Summarization Using Sequence-to-Sequence RNNs and Beyond, August 2016. URL <http://arxiv.org/abs/1602.06023>. arXiv:1602.06023 [cs]. 3
- Shashi Narayan, Shay B. Cohen, and Mirella Lapata. Don't Give Me the Details, Just the Summary! Topic-Aware Convolutional Neural Networks for Extreme Summarization, August 2018. URL <http://arxiv.org/abs/1808.08745>. arXiv:1808.08745 [cs] version: 1. 8

## BIBLIOGRAPHY

- Jun-Ping Ng and Viktoria Abrecht. Better Summarization Evaluation with Word Embeddings for ROUGE, August 2015. URL <http://arxiv.org/abs/1508.06034>. arXiv:1508.06034 [cs]. 26
- Long Ouyang, Jeff Wu, Xu Jiang, Diogo Almeida, Carroll L. Wainwright, Pamela Mishkin, Chong Zhang, Sandhini Agarwal, Katarina Slama, Alex Ray, John Schulman, Jacob Hilton, Fraser Kelton, Luke Miller, Maddie Simens, Amanda Askell, Peter Welinder, Paul Christiano, Jan Leike, and Ryan Lowe. Training language models to follow instructions with human feedback, March 2022. URL <http://arxiv.org/abs/2203.02155>. arXiv:2203.02155 [cs]. 24
- Alec Radford, Karthik Narasimhan, Tim Salimans, and Ilya Sutskever. Improving Language Understanding by Generative Pre-Training. page 12, 2018. 5
- Colin Raffel, Noam Shazeer, Adam Roberts, Katherine Lee, Sharan Narang, Michael Matena, Yanqi Zhou, Wei Li, and Peter J Liu. Exploring the Limits of Transfer Learning with a Unified Text-to-Text Transformer. page 67, 2019. 2, 5, 10, 22
- Alexander M. Rush, Sumit Chopra, and Jason Weston. A Neural Attention Model for Abstractive Sentence Summarization. In *Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing*, pages 379–389, Lisbon, Portugal, September 2015. Association for Computational Linguistics. doi: 10.18653/v1/D15-1044. URL <https://aclanthology.org/D15-1044>. 3
- Rico Sennrich, Barry Haddow, and Alexandra Birch. Improving Neural Machine Translation Models with Monolingual Data, June 2016. URL <http://arxiv.org/abs/1511.06709>. arXiv:1511.06709 [cs]. 2, 6
- Noam Shazeer and Mitchell Stern. Adafactor: Adaptive Learning Rates with Sublinear Memory Cost. In *Proceedings of the 35th International Conference on Machine Learning*, pages 4596–4604. PMLR, July 2018. URL <https://proceedings.mlr.press/v80/shazeer18a.html>. ISSN: 2640-3498. 12
- Vésteinn Snæbjarnarson and Hafsteinn Einarsson. Cross-Lingual QA as a Stepping Stone for Monolingual Open QA in Icelandic. In *Proceedings of the Workshop on Multilingual Information Access (MIA)*, pages 29–36, Seattle, USA, July 2022a. Association for Computational Linguistics. doi: 10.18653/v1/2022.mia-1.4. URL <https://aclanthology.org/2022.mia-1.4>. 6, 23
- Vésteinn Snæbjarnarson and Hafsteinn Einarsson. Natural Questions in Icelandic. In *Proceedings of the Thirteenth Language Resources and Evaluation Conference*, pages 4488–4496, Marseille, France, June 2022b. European Language Resources Association. URL <https://aclanthology.org/2022.lrec-1.477>. 6
- Vésteinn Snæbjarnarson, Haukur Barri Símonarson, Pétur Orri Ragnarsson, Svanhvít Lilja Ingólfssdóttir, Haukur Páll Jónsson, Vilhjálmur Þorsteinsson, and Hafsteinn Einarsson. A Warm Start and a Clean Crawled Corpus – A Recipe for Good

- Language Models, January 2022. URL <http://arxiv.org/abs/2201.05601>. arXiv:2201.05601 [cs]. 6, 23
- Kaitao Song, Xu Tan, Tao Qin, Jianfeng Lu, and Tie-Yan Liu. MASS: Masked Sequence to Sequence Pre-training for Language Generation. *arXiv:1905.02450 [cs]*, June 2019. URL <http://arxiv.org/abs/1905.02450>. arXiv: 1905.02450. 5
- Steinþór Steingrímsson, Sigrún Helgadóttir, Eiríkur Rögnvaldsson, Starkaður Barkarson, and Jón Guðnason. Risamálheild: A Very Large Icelandic Text Corpus. In *Proceedings of the Eleventh International Conference on Language Resources and Evaluation (LREC 2018)*, Miyazaki, Japan, May 2018. European Language Resources Association (ELRA). URL <https://aclanthology.org/L18-1690>. 7
- Ilya Sutskever, Oriol Vinyals, and Quoc V Le. Sequence to Sequence Learning with Neural Networks. In *Advances in Neural Information Processing Systems*, volume 27. Curran Associates, Inc., 2014. URL <https://proceedings.neurips.cc/paper/2014/hash/a14ac55a4f27472c5d894ec1c3c743d2-Abstract.html>. 5
- Haukur Barri Símonarson, Vésteinn Snæbjarnarson, Pétur Orri Ragnarsson, Haukur Páll Jónsson, and Vilhjálmur Þorsteinsson. Miðeind’s WMT 2021 submission. Technical report, September 2021. URL <https://ui.adsabs.harvard.edu/abs/2021arXiv210907343B>. Publication Title: arXiv e-prints ADS Bibcode: 2021arXiv210907343B Type: article. 6
- Martin A. Tanner and Wing Hung Wong. The Calculation of Posterior Distributions by Data Augmentation. *Journal of the American Statistical Association*, 82(398):528–540, June 1987. ISSN 0162-1459. doi: 10.1080/01621459.1987.10478458. URL <https://www.tandfonline.com/doi/abs/10.1080/01621459.1987.10478458>. Publisher: Taylor & Francis \_eprint: <https://www.tandfonline.com/doi/pdf/10.1080/01621459.1987.10478458>. 2
- Oguzhan Tas and Farzad Kiyani. A SURVEY AUTOMATIC TEXT SUMMARIZATION. *PressAcademia Procedia*, 5(1):205–213, June 2017. doi: 10.17261/Pressacademia.2017.591. URL <https://dergi.park.org.tr/en/pub/pap/issue/33433/371658>. Number: 1. 1
- Chau Tran, Shruti Bhosale, James Cross, Philipp Koehn, Sergey Edunov, and Angela Fan. Facebook AI WMT21 News Translation Task Submission. *arXiv:2108.03265 [cs]*, August 2021. URL <http://arxiv.org/abs/2108.03265>. arXiv: 2108.03265. 9
- Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N Gomez, Łukasz Kaiser, and Illia Polosukhin. Attention is All you Need. In



## BIBLIOGRAPHY

- Advances in Neural Information Processing Systems*, volume 30. Curran Associates, Inc., 2017. URL <https://proceedings.neurips.cc/paper/2017/hash/3f5ee243547dee91fbd053c1c4a845aa-Abstract.html>. 1, 3
- Linting Xue, Noah Constant, Adam Roberts, Mihir Kale, Rami Al-Rfou, Aditya Siddhant, Aditya Barua, and Colin Raffel. mT5: A massively multilingual pre-trained text-to-text transformer, March 2021. URL <http://arxiv.org/abs/2010.11934>. arXiv:2010.11934 [cs]. 6, 10
- Manzil Zaheer, Guru Guruganesh, Avinava Dubey, Joshua Ainslie, Chris Alberti, Santiago Ontanon, Philip Pham, Anirudh Ravula, Qifan Wang, Li Yang, and Amr Ahmed. Big Bird: Transformers for Longer Sequences. page 15, 2020. 4
- Jingqing Zhang, Yao Zhao, Mohammad Saleh, and Peter Liu. PEGASUS: Pre-training with Extracted Gap-sentences for Abstractive Summarization. In *International Conference on Machine Learning*, pages 11328–11339. PMLR, November 2020. URL <https://proceedings.mlr.press/v119/zhang20ae.html>. ISSN: 2640-3498. 2, 5, 9
- Fuzhen Zhuang, Zhiyuan Qi, Keyu Duan, Dongbo Xi, Yongchun Zhu, Hengshu Zhu, Hui Xiong, and Qing He. A Comprehensive Survey on Transfer Learning. *Proceedings of the IEEE*, 109(1):43–76, January 2021. ISSN 1558-2256. doi: 10.1109/JPROC.2020.3004555. Conference Name: Proceedings of the IEEE. 2
- Barret Zoph, Deniz Yuret, Jonathan May, and Kevin Knight. Transfer Learning for Low-Resource Neural Machine Translation, April 2016. URL <http://arxiv.org/abs/1604.02201>. arXiv:1604.02201 [cs]. 1

# A. Translated Datasets' Examples

Examples of the Icelandic language translations of Xsum and CNN/DailyMail, obtained with the Facebook WMT21 model, as well as their original English documents.

## A.1. XSum

English	
<b>Document</b>	The 24-year-old Belgium-born midfielder has made just 23 Premier League starts since joining from Porto on a five-and-a-half-year contract in February 2016. "Sometimes things just don't pan out as you hope," Hughes said. "It has been a frustration for all parties involved. It's fair to say with Giannelli that things haven't quite worked out." Imbula became Portuguese football's record signing when he joined Porto in a £15.1m deal from Marseille in July 2015. "He hasn't quite adjusted to the Premier League," added Hughes. "If there is interest in him then we would consider that. We haven't had any interest, in fairness, but we would look at things if the right offer came in."
<b>Summary</b>	Stoke City's £18.3m record signing Giannelli Imbula could leave the club this summer, manager Mark Hughes says.
Icelandic Translation	
<b>Document</b>	Þessi 24 ára gamli miðjumaður sem er fæddur í Belgíu hefur aðeins byrjað 23 leiki í ensku úrvalsdeildinni síðan hann kom frá Porto á fimm og hálfis árs samningi í febrúar 2016. Stundum ganga hlutirnir bara ekki upp eins og þú vonast til sagði Hughes. "Þetta hefur verið svekkjandi fyrir alla aðila. Það er sanngjarnt að segja með Giannelli að hlutirnir hafa ekki alveg gengið upp". Imbula varð metkaup portúgalska fótoltans þegar hann gekk til liðs við Porto á 15,1 milljón punda frá Marseille í júlí 2015. " Hann hefur ekki alveg aðlagast ensku úrvalsdeildinni " bætti Hughes við. "Ef það er áhugi á honum þá myndum við fluga það. Við höfum ekki haft neinn áhuga, í sanngirni sagt, en við myndum skoða hlutina ef rétta tilboðið kæmi".
<b>Summary</b>	Mark Hughes knattspyrnustjóri Stoke City segir að Giannelli Imbula sem félagið keypti á 18,3 milljónir punda gæti yfirgefið félagið í sumar

English	
<b>Document</b>	<p>In the app, the cartoon characters feature with the real world in the background. The developers claim it is the first time in the UK that "augmented reality" has been used in this way. They and Conwy council hope it will bring tourists to the town where the inspiration for Alice went on holiday. The original launch of the app in August was delayed after key figures for tourists to follow were altered. Virtual characters failed to recognise sculptures in the town after Conwy council painted them. Next year marks 150 years since Alice in Wonderland was first published. Simon Burrows, co-director of Alice in Wonderland Limited which has co-ordinated the project, said: "People who use the app can look at Llandudno landmarks on screen, while characters from the Alice books tell the story behind them. "It's also possible to take a picture of yourself or your friends standing next to the Mad Hatter, the Cheshire Cat or Alice herself." The Alice in Wonderland story, and its sequel Through the Looking Glass, were written by Lewis Carroll (whose real name was Charles Dodgson). He composed the original story to entertain a young girl called Alice Liddell, with whose family he was friends. Alice Liddell's family lived in Oxford but had a house in Llandudno and spent many of their holidays there. Company director Barry Mortlock said: "I've lived in Llandudno for 36 years now, and it worried me that the town wasn't doing enough in recent years to celebrate its links with the real life Alice. "The app is now ready for next year's important anniversary. "It has the potential to attract more visitors from overseas, where there's a huge interest in the Alice in Wonderland stories." Graham Rees, Conwy council's cabinet member for tourism, said: "This will be a huge asset to Llandudno and really enhance what we can offer to visitors."</p>
<b>Summary</b>	<p>Characters from the Alice in Wonderland books will guide visitors on a tour of Llandudno in a new app which shows the town's links with the stories.</p>
Icelandic Translation	
<b>Document</b>	<p>Í forritinu eru teiknimyndapersónurnar með raunheiminn í bakgrunni. Framleiðendurnir halda því fram að þetta sé í fyrsta sinn í Bretlandi sem "viðbættur veruleiki" e. augmented reality er notaður á þennan hátt. Þau og bæjarstjórn Conwy vonast til þess að þetta verði til þess að ferðamenn flykkist til bæjarins þar sem innblásturinn að Lísu fór í frí. Upphaflegri útgáfu smáforritsins sem átti að koma út í ágúst var frestað eftir að lykiltölum fyrir ferðamenn var breytt. Sýndarpersónum mistókst að bera kennsl á höggmyndir í bænum eftir að bæjarstjórn Conwy málaði þær. Á næsta ári verða liðin 150 ár frá því að Lísa í Undralandi kom fyrst út. Simon Burrows meðstjórnandi Alice in Wonderland Limited sem hefur samhæft verkefnið sagði "Fólk sem notar appið getur skoðað kennileiti Llandudno á skjánum á meðan persónur úr bókunum um Lísu segja söguna á bak við þau. "Það er líka hægt að taka mynd af þér eða vinum þínum standandi við hliðina á Brjálaða hattaranum, Cheshire kettinum eða Lísu sjálfri". Sagan um Lísu í Undralandi og framhald hennar, Í gegnum spegilinn, var skrifuð af Lewis Carroll (sem hét réttu nafni Charles Dodgson). Hann samdi upprunalegu söguna til að skemmta ungri stúlku að nafni Alice Liddell en hann var vinur fjölskyldu hennar. Fjölskylda Alice Liddell bjó í Oxford en átti hús í Llandudno og eyddi mörgum af fríum sínum þar. Leikstjóri myndarinnar Barry Mortlock sagði Ég hef búið í Llandudno í 36 ár og það olli mér áhyggjum að bærinn hefði ekki gert nóg á undanförunum árum til að halda upp á tengsl sín við hina raunverulegu Lísu. "Appið er nú tilbúið fyrir mikilvæga afmælishátíð á næsta ári. "Það hefur möguleika á að laða að fleiri gesti frá útlöndum, þar sem það er gríðarlegur áhugi á sögunum um Lísu í Undralandi". Graham Rees, ferðamálafulltrúi Conwy-ráðsins, sagði: "Þetta verður mikill fengur fyrir Llandudno og eykur verulega það sem við getum boðið gestum".</p>
<b>Summary</b>	<p>Stafir úr Lísa í Undralandi bókum mun leiða gesti í skoðunarferð um Llandudno í nýju forriti sem sýnir tengla bæjarins með sögunum.</p>

---

English

---

<b>Document</b>	Operators such as Uber and GrabTaxi will now have to register with the country's Land Transport Authority. Under the regulations, fees charged by the booking apps cannot exceed those of regular taxi companies. The new measures will be enforced from the second quarter of next year, the government office said on Friday. "To ensure that taxi services remain equally accessible to all members of the public, bidding and pre-trip tipping for taxi services will not be allowed," the LTA added in a statement. Wealthy city-state Singapore is among the most expensive places in the world to own a car, with many commuters relying on taxis to get around. The popularity of taxi booking services has surged in the city of nearly 5.4 million people, especially during peak business hours. Uber has expanded rapidly in Singapore offering services such as high-end UberExec and cheaper UberX cars, along with a service to book regular taxis. The company said the new rules would not affect its extra services, because it already has partnerships with licensed limousine and rental car companies. "We appreciate that the LTA has acknowledged the benefits our technology brings, and like Uber, is putting the interest and safety of consumers and drivers first," said Uber's regional general manager of Southeast Asia, Mike Brown, in a statement to Reuters. The San Francisco-based company has faced protests around the world by taxi drivers who are against its services, which allow people to become unlicensed taxi drivers by using their own cars.
<b>Summary</b>	Singapore plans to regulate third-party taxi booking apps such as Uber by capping fees and limiting them to use only licensed vehicles and drivers.

---

## Icelandic Translation

<b>Document</b>	<p>Rekstraraðilar eins og Uber og GrabTaxi verða nú að skrá sig hjá Landflutningastofnun landsins. Samkvæmt reglunum mega gjöldin sem bóknaröppin innheimta ekki vera hærri en hjá venjulegum leigubílafyrirtækjum. Nýju aðgerðirnar verða innleiddar frá og með öðrum ársfjórðungi næsta árs að því er skrifstofa ríkisstjórnarinnar greindi frá á föstudag. "Til að tryggja að leigubílaþjónusta verði jafn aðgengileg öllum almenningi verður ekki leyfilegt að bjóða í leigubílaþjónustu eða gefa þjóf fé fyrir akstur" sagði í yfirlýsingu frá LTA. Auðuga borgríkið Singapur er á meðal dýrustu staða í heimi til að eiga bíl og margir borgarbúar reiða sig á leigubíla til að komast á milli staða. Vinsældir bóknarþjónustu fyrir leigubíla hafa stóraukist í borginni þar sem búa tæplega 5,4 milljónir manna sérstaklega á háannatímum. Uber hefur stækkað hratt í Singapur og býður upp á þjónustu á borð við hágæða UberExec og ódýrari UberX bíla ásamt þjónustu til að bóka venjulega leigubíla. Fyrirtækið sagði að nýju reglurnar myndu ekki hafa áhrif á aukþjónustu þess vegna þess að það á nú þegar í samstarfi við fyrirtæki sem hafa leyfi til að leigja lúxusbifreiðar og bílaleigubíla. "Við kunnum að meta að LTA hefur viðurkennt þann ávinning sem tæknin okkar hefur í för með sér og líkt og Uber setur það hagsmuni og öryggi neytenda og ökumanna í fyrsta sæti" sagði Mike Brown svæðisstjóri Uber í Suðaustur-Asíu í yfirlýsingu til Reuters. Fyrirtækið sem er með höfuðstöðvar í San Francisco hefur mætt mótmælum um allan heim frá leigubílstjórum sem eru á móti þjónustu þess sem gerir fólki kleift að gerast leyfislausir leigubílstjórar með því að nota eigin bíla</p>
<b>Summary</b>	Stjórnvöld í Singapur hyggjast setja hömlur á leigubílaþjónustur á borð við Uber með því að setja þak á gjaldtöku og takmarka notkun þeirra við ökutæki og bílstjóra með tilskilin leyfi

---

English	
<b>Document</b>	In the app, the cartoon characters feature with the real world in the background. The developers claim it is the first time in the UK that "augmented reality" has been used in this way. They and Conwy council hope it will bring tourists to the town where the inspiration for Alice went on holiday. The original launch of the app in August was delayed after key figures for tourists to follow were altered. Virtual characters failed to recognise sculptures in the town after Conwy council painted them. Next year marks 150 years since Alice in Wonderland was first published. Simon Burrows, co-director of Alice in Wonderland Limited which has co-ordinated the project, said: "People who use the app can look at Llandudno landmarks on screen, while characters from the Alice books tell the story behind them. "It's also possible to take a picture of yourself or your friends standing next to the Mad Hatter, the Cheshire Cat or Alice herself." The Alice in Wonderland story, and its sequel Through the Looking Glass, were written by Lewis Carroll (whose real name was Charles Dodgson). He composed the original story to entertain a young girl called Alice Liddell, with whose family he was friends. Alice Liddell's family lived in Oxford but had a house in Llandudno and spent many of their holidays there. Company director Barry Mortlock said: "I've lived in Llandudno for 36 years now, and it worried me that the town wasn't doing enough in recent years to celebrate its links with the real life Alice. "The app is now ready for next year's important anniversary. "It has the potential to attract more visitors from overseas, where there's a huge interest in the Alice in Wonderland stories." Graham Rees, Conwy council's cabinet member for tourism, said: "This will be a huge asset to Llandudno and really enhance what we can offer to visitors."
<b>Summary</b>	Characters from the Alice in Wonderland books will guide visitors on a tour of Llandudno in a new app which shows the town's links with the stories.
Icelandic Translation	
<b>Document</b>	Í forritinu eru teiknimyndapersónurnar með raunheiminn í bakgrunni. Framleiðendurnir halda því fram að þetta sé í fyrsta sinn í Bretlandi sem "viðbættur veruleiki" e. augmented reality er notaður á þennan hátt. Þau og bæjarstjórn Conwy vonast til þess að þetta verði til þess að ferðamenn flykkist til bæjarins þar sem innblásturinn að Lísu fór í frí. Upphaflegri útgáfu smáforritsins sem átti að koma út í ágúst var frestað eftir að lykiltölum fyrir ferðamenn var breytt. Sýndarpersónum mistókst að bera kennsl á höggmyndir í bænum eftir að bæjarstjórn Conwy málaði þær. Á næsta ári verða liðin 150 ár frá því að Lísa í Undralandi kom fyrst út. Simon Burrows meðstjórnandi Alice in Wonderland Limited sem hefur samhæft verkefnið sagði "Fólk sem notar appið getur skoðað kennileiti Llandudno á skjánum á meðan persónur úr bókunum um Lísu segja söguna á bak við þau. "Það er líka hægt að taka mynd af þér eða vinum þínum standandi við hliðina á Brjálaða hattaranum, Cheshire kettinum eða Lísu sjálfri". Sagan um Lísu í Undralandi og framhald hennar, Í gegnum spegilinn, var skrifuð af Lewis Carroll (sem hét réttu nafni Charles Dodgson). Hann samdi upprunalegu söguna til að skemmta ungri stúlku að nafni Alice Liddell en hann var vinur fjölskyldu hennar. Fjölskylda Alice Liddell bjó í Oxford en átti hús í Llandudno og eyddi mörgum af fríum sínum þar. Leikstjóri myndarinnar Barry Mortlock sagði Ég hef búið í Llandudno í 36 ár og það olli mér áhyggjum að bærinn hefði ekki gert nóg á undanförunum árum til að halda upp á tengsl sín við hina raunverulegu Lísu. "Appið er nú tilbúið fyrir mikilvæga afmælishátíð á næsta ári. "Það hefur möguleika á að laða að fleiri gesti frá útlöndum, þar sem það er gríðarlegur áhugi á sögunum um Lísu í Undralandi". Graham Rees, ferðamálafulltrúi Conwy-ráðsins, sagði: "Þetta verður mikill fengur fyrir Llandudno og eykur verulega það sem við getum boðið gestum".
<b>Summary</b>	Stafir úr Lísa í Undralandi bókum mun leiða gesti í skoðunarferð um Llandudno í nýju forriti sem sýnir tengla bæjarins með sögunum.

---

English

---

<b>Document</b>	Of these 216 planets, 10 of them could potentially be home to some sort of life, because they are 'Earth-like'. This means they have 'rocky' characteristics, and they are located in the 'Goldilocks Zone' of their solar system. If a planet is in a 'Goldilocks Zone' it is not too hot or too cold, which means water can be in its liquid form there. The Kepler Telescope has made some amazing discoveries since it was launched in 2009, but here are the top five findings... In 2014 the telescope made a huge discovery by finding a planet nicknamed: 'Earth 2.0'. The planet called, Kepler-452b has been dubbed as 'Earth 2.0', because it shares many characteristics with our Earth, even though it is 1,400 light years away from us. The first potentially habitable planet the telescope found was Kepler-22b, way back in 2011. Scientists agreed the planet was could be liveable because it was the perfect distance from its sun. In April 2017 the telescope found a super-Earth. This is a planet called LHS 1140b. It is a good candidate for life and orbits a red dwarf star 40 light years away. However, scientists think the planet might have had an ocean made entirely out of magma! - which released steam into the atmosphere. In February 2017, Kepler discovered seven new Earth-sized exoplanets. Exoplanets are planets that orbit a different star to our Sun - in a different solar system to the one we are in. In 2011 Nasa discovered the first planet to orbit two suns! It is called Kepler-16b, and is roughly 200 light years away from Earth. Some scientists have compared it to Luke Skywalker's home planet Tatooine, from the Star Wars movies, which also had two suns.
<b>Summary</b>	Nasa's Kepler telescope has recently found 216 new planets, meaning it has now discovered more than 4000 potential new worlds!

---

## Icelandic Translation

<b>Document</b>	Af þessum 216 plánetum gætu 10 þeirra hugsanlega verið heimili einhvers konar lífs vegna þess að þau eru "jörð eins og". Það þýðir að þeir hafa "grjótharða" eiginleika og þeir eru staðsettir á "Gullbrásvæðinu" í sínu sólkerfi. Ef pláneta er á svokölluðu Gullbrásvæði er hvorki of heitt né of kalt sem þýðir að vatn getur verið á fljótandi formi þar. Kepler-sjónaukinn hefur gert nokkrar stórkostlegar uppgötvanir síðan honum var skotið á loft árið 2009 en hér eru fimm helstu uppgötvánirnar. Árið 2014 gerði sjónaukinn mikla uppgötvun með því að finna plánetu sem var kölluð "Jörð 2.0". Reikistjarnan sem kallast Kepler-452b hefur verið kölluð "Jörð 2.0" þar sem hún deilir mörgum einkennum með Jörðinni, þrátt fyrir að vera í 1.400 ljósára fjarlægð frá okkur. Fyrsta mögulega lífvænlega plánetan sem sjónaukinn fann var Kepler-22b, árið 2011. Vísindamenn voru sammála um að plánetan væri lífvænleg vegna þess að hún væri í fullkominni fjarlægð frá sólinni. Í apríl 2017 fann sjónaukinn risajörð. Þetta er reikistjarna sem kallast LHS 1140b. Hún er góður kandiðat fyrir líf og gengur um rauða dvergstjörnu í 40 ljósára fjarlægð. Vísindamenn telja hins vegar að plánetan gæti hafa haft haf sem var eingöngu gert úr kviku - sem losaði gufu út í andrúmsloftið. Í febrúar 2017 uppgötvaði Kepler sjö nýjar fjarreikistjörnur á stærð við jörðina. Fjarreikistjörnur eru reikistjörnur sem ganga um aðra stjörnu en sólina okkar - í öðru sólkerfi en því sem við búum í. Árið 2011 fann NASA fyrstu plánetuna sem hringsólar um tvær sólar! Hún kallast Kepler-16b og er í um 200 ljósára fjarlægð frá jörðinni. Sumir vísindamenn hafa líkt henni við Tatooine heimaplánetu Luke Skywalker úr Stjörnustríðsmyndunum sem einnig hafði tvær sólar
<b>Summary</b>	Kepler sjónauki NASA hefur nýlega fundið 216 nýjar plánetur, sem þýðir að það hefur nú uppgötvað meira en 4000 hugsanlega nýja heima!

---

## A.2. CNN/DailyMail

English	
<b>Document</b>	A bobcat was snapped hauling a shark along a Florida beach in an unbelievable photograph. The image was taken at Sebastian Inlet State Park near Vero Beach by John Bailey, WPTV reported. Bailey told the television station that the bobcat collected the shark after entering the water. Scroll down for video . Life in the water: A bobcat was snapped hauling a shark along Sebastian Inlet State Park . However, Bailey told WPTV the bobcat might have been scared by his presence, since it scampered off toward the woods and let go of its find. Amy Kight, Busch Wildlife Sanctuary's Animal Care Director, told WPTV: 'My first reaction was amazement that somebody actually took that picture because bobcats are generally very, very reclusive. 'You don't really see them out and about too often.' A spokesman for the Florida Fish and Wildlife Conservation Commission has said she believes the image is genuine. FWC's Liz Barraco told WESH 'There is no reason to believe it's fake.' The bobcat exhibited typical behavior, FWC officials told the NBC affiliate station. Scene: The image was reportedly taken at Sebastian Inlet State Park (pictured) near Vero Beach, Florida .
<b>Summary</b>	The image was taken at Sebastian Inlet State Park near Vero Beach, Florida, by John Bailey . Bailey has said the bobcat caught the shark after entering the water . The bobcat might have been scared by his presence, since it scampered off toward the woods and let go of its find . A spokesman for the Florida Fish and Wildlife Conservation Commission believes the image is genuine .
Icelandic Translation	
<b>Document</b>	Á ótrúlegri mynd náðist af rauðbrystingi draga hákarl eftir strönd í Flórída. Myndin var tekin í Sebastian Inlet State Park nálægt Vero Beach af John Bailey að sögn WPTV. Bailey sagði í samtali við sjónvarpsstöðina að rauðbrystingurinn hefði safnað hákarlinum saman eftir að hann fór í vatnið. Skrolla niður fyrir vídeó. Líf í vatni: Bobcat var smellt á að draga hákarl meðfram Sebastian Inlet State Park. Bailey sagði hins vegar í samtali við WPTV að bobcatinn gæti hafa verið hræddur við nærveru hans þar sem hann hljóp í átt að skóginum og sleppti því sem hann fann. Amy Kight yfirmaður dýragæslu Busch Wildlife Sanctuary sagði í samtali við WPTV " Mín fyrstu viðbrögð voru undrun yfir því að einhver hafi í raun tekið þessa mynd því að rauðbrystingar eru yfirleitt mjög einræn dýr. "Maður sér þá ekki oft á ferli". Talsmaður dýraverndunarsamtaka í Flórída, Florida Fish and Wildlife Conservation Commission, segist sannfærður um að myndin sé ósvikin.. Liz Barraco hjá FWC sagði í samtali við WESH " Það er engin ástæða til að ætla að þetta sé falsað " Kötturinn sýndi dæmigerða hegðun að sögn starfsmanna FWC. Atriði: Myndin var að sögn tekin í Sebastian Inlet State Park (sjá mynd) nálægt Vero Beach, Flórída.
<b>Summary</b>	Myndin var tekin í Sebastian Inlet State Park nálægt Vero Beach, Flórída, af John Bailey. Bailey hefur sagt að Bobcat caught hákarl eftir að slá vatn. The Bobcat gæti hafa verið hrædd við nærveru hans, þar sem það scampered burt í átt að skóginum og láta fara af finna sína. Talsmaður Florida Fish and Wildlife Conservation framkvæmdastjórnarinnar telur að myndin er ósvikinn.

---

English

---

<b>Document</b>	This is the moment a clumsy baby elephant took a tumble as he rushed to catch up with his mother as she crossed a road. The cute calf had been taking a stroll through the bush with his mother at Idube Game Reserve in South Africa. When the pair came across a dirt track road, the adult strolled on ahead confidently. Careful! A clumsy elephant takes a tumble as he rushes to catch up with his mother as she crosses the road . But when the baby elephant rushed to catch her up he appeared to get his leg caught in the grass. The clumsy little one stumbled as he stepped down onto the path and struggled to regain his balance. Eventually he fell over and landed - trunk first - down in the mud. The cute calf had been taking a stroll through the bush with his mother at Idube Game Reserve in South Africa . But when the baby elephant rushed to catch her up he appeared to get his leg caught in the grass . The clumsy little one stumbled as he stepped down onto the path, eventually losing his balance altogether . Faceplant: The young elephant eventually he fell over and landed - trunk first - down in the mud . Luckily the young elephant did not appear to hurt anything but his pride and was quickly back on his feet. Noticing her clumsy calf's fall, the mother elephant, who was carrying a huge tree branch in her trunk, had halted her onward march and turned back to check her baby. Having given him a once over, the pair disappeared back into the bush. The incredible moment was captured by reserve ranger and wildlife filmmaker Rob Vamplew, known as Rob the Ranger. Luckily he did not appear to hurt anything but his pride and was quickly back on his feet . Noticing her clumsy calf's fall, the mother elephant, who was carrying a huge tree branch in her trunk, had halted her onward march . The giant African elephant went back to her baby to make sure he was ok after his trip .
<b>Summary</b>	Cute but clumsy elephant calf was following his mother though the bush . Suddenly the youngster appeared to catch his foot and his lose balance . Fell over and landed - trunk first - down in the mud at Idube Game Reserve .

---

## Icelandic Translation

<b>Document</b>	Petta er augnablikið þegar klaufalegur fílsungi datt þegar hann var að flýta sér að ná móður sinni þegar hún var að fara yfir götu. Kálfurinn krúttlegi hafði verið á rölti í skóginum með móður sinni á Idube Game Reserve í Suður-Afríku. Þegar parið kom yfir moldarveg rölti sá fullorðni öruggur áfram. Gættu þín! Klaufalegur fill dettur þegar hann flýtir sér að ná móður sinni þegar hún er að fara yfir götuna. En þegar fillinn hljóp til að ná henni upp virtist hann festa fótinn í grasinu. Litli klaufinn hrasaði þegar hann steig niður á stíginn og átti í erfiðleikum með að ná jafnvæginu aftur. Að lokum féll hann um koll og lenti - skottinu á undan - ofan í leðjunni. Kálfurinn krúttlegi hafði verið á rölti í skóginum með móður sinni á Idube Game Reserve í Suður-Afríku. En þegar fillinn hljóp til að ná henni upp virtist hann festa fótinn í grasinu. Litli klaufinn hrasaði þegar hann steig niður á stíginn og missti að lokum jafnvægið. Andlitsplantan: Ungi fillinn féll að lokum um koll og lenti - skottinu á undan - í leðjunni. Til allrar hamingju virtist ungi fillinn ekki meiða neitt nema stolt sitt og var fljótt kominn aftur á fæturna. Þegar hún tók eftir falli kálfsins hafði móðurfillinn sem hélt á stórri trjágrein í rananum stöðvað göngu sína og snúið við til að huga að barninu. Eftir að hafa gefið honum á kjaftinn einu sinni enn hvarf parið aftur inn í kjarrið. Þetta ótrúlega augnablik náðist á myndband af landverði og dýralífskvikmyndagerðarmanninum Rob Vamplew sem gengur undir nafninu Rob the Ranger. Til allrar hamingju virtist hann ekki hafa meitt neitt nema stolt sitt og var fljótur á fætur aftur. Eftir að hafa tekið eftir falli kálfsins hafði móðurfillinn sem hélt á stórri trjágrein í rananum stöðvað göngu hans. Risastóri afríski fillinn fór aftur til barnsins síns til að ganga úr skugga um að það væri í lagi með hann eftir ferðalagið.
<b>Summary</b>	Sætur en klaufalegt fil kálfur var eftir móður sinni þó Bush. Skyndilega unglingur virtist grípa fótinn og missa jafnvægi hans. Féll yfir og lenti - skottinu fyrst - niður í drullu á Idube Game Reserve.

---



## A. Translated Datasets' Examples

---

English	
<b>Document</b>	Nedum Onuoha is wanted by Hull City. The QPR centre-back is one of several players poised to leave should the Loftus Road club go down as they consider streamlining their squad. The 28-year-old has 15 months left on contract and has played regularly under Chris Ramsey. Onuoha was in the starting eleven for the Saturday's crucial 4-1 win over relegation rivals West Bromwich Albion at the Hawthorns. Nedum Onuoha is being targeted by Hull City but is also interesting Everton and West Ham United . Stoke manager Mark Hughes who signed him for QPR is also interested as are Everton and West Ham. QPR have held talks over a move for Catania defender Nicolas Spolli. The 32-year-old has been on loan at Roma. Hull, meanwhile, are also showing interest in West Bromwich Albion striker Victor Anichebe who is likely to be sold this summer. Victor Anichebe (centre) is also being watched by Hull City and could move to the KC Stadium in the summer .
<b>Summary</b>	Hull City are interested in signing Queens Park Rangers' Nedum Onuoha . Everton and West Ham United are also keen on the versatile defender . QPR are in talks with Catania over a deal to sign defender Nicolas Spolli . Hull also have their eye on West Brom attacker Victor Anichebe .

---

Icelandic Translation	
<b>Document</b>	Nedum Onuoha er eftirsóttur af Hull City. Miðvörður QPR er einn af nokkrum leikmönnum sem eru tilbúnir að fara ef félagið á Loftus Road fer niður þar sem þeir íhuga að hagræða í leikmannahópnum. Þessi 28 ára gamli leikmaður á 15 mánuði eftir af samningi sínum og hefur spilað reglulega undir stjórn Chris Ramsey. Onuoha var í byrjunarliðinu í mikilvægum 4 - 1 sigri á fallbaráttuliði West Bromwich Albion á Hawthorns á laugardaginn. Nedum Onuoha er í sigtinu hjá Hull City en er einnig áhugaverður Everton og West Ham United. Mark Hughes knattspyrnustjóri Stoke sem keypti hann til QPR hefur einnig áhuga sem og Everton og West Ham. QPR hefur átt í viðræðum um kaup á Nicolas Spolli varnarmanni Catania. Þessi 32 ára gamli leikmaður hefur verið í láni hjá Roma. Á sama tíma hefur Hull einnig áhuga á Victor Anichebe framherja West Bromwich Albion sem verður líklega seldur í sumar. Victor Anichebe (fyrir miðju) er einnig undir smásjá Hull City og gæti farið á KC leikvanginn í sumar.
<b>Summary</b>	Hull City hefur áhuga á að fá Nedum Onuoha frá Queens Park Rangers Everton og West Ham United hafa einnig áhuga á þessum fjölhæfa varnarmanni QPR er í viðræðum við Catania um kaup á varnarmanninum Nicolas Spolli og Hull hefur einnig augastað á Victor Anichebe sóknarmanni West Brom

---

---

**English**


---

<b>Document</b>	Richard Sherman may have been devastated after he and the Seattle Seahawks fell to the New England Patriots in this year's Super Bowl, but it did not last for long. On February 5, just four days after that agonizing defeat, Sherman and girlfriend Ashley Moss welcomed a son, Rayden. And now, Sherman is sharing photos of his baby boy. Scroll down for video . Richard Sherman shared a photo of him and his girlfriend Ashley Moss with their son Rayden on Instagram (courtesy of estevens) Rayden (above) was born on February 5, just a few days after Sherman and the Seattle Seahawks lost the Super Bowl to the New England Patriots . Sherman congratulates Patriots quarterback Tom Brady (above) after the game . In the most recent photo, which he posted to Instagram on Thursday, Sherman and Moss are seen cradling Rayden, with the football star writing; 'Best part of Life.' He previously shared a photo that showed him holding Rayden's hand shortly after his birth and one of the newborn with his feet in the sand on the beach. And while the highly coveted Vince Lombardi Trophy may not have gone to Sherman and the Seahawks this year, he did manage to snag another major accolade for his work during the 2014 season as the cornerback was named the NFC Defensive Player of the Year by the Kansas City Committee of 101. Sherman and Moss (above) in a photo he shared last month . It was just announced the Seahawks would play the Denver Broncos in their first preseason game . The off-season will not be all fun for Sherman though, as he will also be undergoing surgery to repair a torn ligament in his left elbow. This after he signed a \$57.4 million contract extension last May that will keep him in a Seahawks uniform for at least three more seasons. The 27-year-old will return to the field in August, as it has been announced that the Seahawks will take on the Denver Broncos in their first preseason game as they begin their quest to make it to the Super Bowl for the third consecutive season.
<b>Summary</b>	Richard Sherman shared a photo of his son Rayden on Instagram . Rayden was born on February 5, just a few days after Sherman and the Seattle Seahawks lost the Super Bowl to the New England Patriots . Rayden's mother is Sherman's girlfriend, Ashley Moss . It was just announced the Seahawks would play the Denver Broncos in their first preseason game .

---

**Icelandic Translation**


---

<b>Document</b>	Richard Sherman kann að hafa verið rúst eftir að hann og Seattle Seahawks féll til New England Patriots í Super Bowl á þessu ári, en það var ekki lengi. Þann 5. febrúar aðeins fjórum dögum eftir þennan kvalafulla ósigur eignuðust Sherman og kærasta hans Ashley Moss soninn Rayden. Og nú er Sherman farinn að deila myndum af syni sínum. Skrolla niður fyrir vídeó. Richard Sherman deildi mynd af sér og kærustu sinni Ashley Moss með son þeirra Rayden á Instagram (með leyfi Estevens) Rayden (fyrir ofan) fæddist 5. febrúar, aðeins nokkrum dögum eftir að Sherman og Seattle Seahawks töpuðu Ofurskálinni fyrir New England Patriots. Sherman óskar Tom Brady leikstjórnanda Patriots til hamingju hér að ofan eftir leikinn. Á nýjustu myndinni, sem hann birti á Instagram á fimmtudag, sjást Sherman og Moss halda á Rayden og fótboltastjarnan skrifar: "Besti hluti lífsins". Áður hafði hann deilt mynd sem sýndi hann halda í hönd Rayden skömmu eftir fæðingu hans og einn nýburann með fæturna í sandinum á ströndinni.. Og þrátt fyrir að hinn eftirsótti Vince Lombardi-bikar hafi ekki farið til Sherman og Seahawks í ár náði hann að krækja í aðra stóra viðurkenningu fyrir störf sín á tímabilinu 2014 þegar hornamaðurinn var valinn varnarleikmaður ársins í NFC-deildinni af nefnd 101 frá Kansas City. Sherman og Moss (fyrir ofan) á mynd sem hann deildi í síðasta mánuði. Nýlega var tilkynnt að Seahawks myndu mæta Denver Broncos í fyrsta undirbúningsleik liðsins. Undirbúningstímabilið verður þó ekki eintóm skemmtun fyrir Sherman því hann mun einnig gangast undir aðgerð til að laga rifid liðband í vinstri ölboga. Þetta eftir að hann skrifaði undir 57,4 milljón dollara samning í maí síðastliðnum sem mun halda honum í búningi Seahawks í að minnsta kosti þrjú tímabil í viðbót. Þessi 27 ára gamli leikmaður mun snúa aftur á völlinn í ágúst en tilkynnt hefur verið að Seahawks muni mæta Denver Broncos í fyrsta undirbúningsleik liðsins fyrir tímabilið þar sem liðið stefnir að því að komast í Super Bowl þriðja tímabilið í röð
<b>Summary</b>	Richard Sherman deildi mynd af syni sínum Rayden á Instagram. Rayden fæddist 5. febrúar, aðeins nokkrum dögum eftir að Sherman og Seattle Seahawks töpuðu Super Bowl fyrir New England Patriots. Móðir Rayden er kærasta Sherman, Ashley Moss. Það var bara tilkynnt að Seahawks myndi spila Denver Broncos í fyrsta leik sínum á undirbúningstímabilinu.

---

## A. Translated Datasets' Examples

English	
<b>Document</b>	<p>A New York detective has been suspended after being accused of robbing a store while colleagues were arresting workers for selling illegal cigarettes. The manager of the deli in Brooklyn said Detective Ian Cyrus, who has been suspended without pay, was caught on camera stealing cash. His supervisor has been placed on desk duty. Scroll down for video . Caught on camera: On viewing the CCTV security cameras, the store manager said he saw one of the detectives finding the rent money box under the counter, and Detective Ian Cyrus grabbing a handful of it . Mr Abdullah said Detective Ian Cyrus then bends down and when he stands up, he appears to put the money in his coat pocket . Five New York Police Department (NYPD) detectives went to the Brooklyn store on Friday April 3 and arrested two employees accused of selling untaxed cigarettes, according to ABC7 Eyewitness News. Officers reportedly took money from the cash register and packs of cigarettes as evidence. Manager Ali Abdullah, who was not in the shop at the time, was given a receipt from the police the following day saying \$593 had been seized, but he realised \$2,650 in store rent money was missing. On viewing the CCTV security cameras, he said Detective Cyrus grabbed a handful of cash from the rent money box under the counter. Mr Abdullah said the detective then bends down and when he stands up, he appears to put the money in his coat pocket. 'I look at my system, I see the officer took the money,' Mr Abdullah told the news channel. 'It's crazy.' He said he called a nearby police department, who sent a supervisor to look at the video. Mr Abdullah claims the supervisor said: 'He's going to lose his pension, lose his life, for \$2,600?' NYPD Deputy Commissioner Stephen Davis told MailOnline: 'Based on the nature of the allegations in this incident, in addition to the video provided to us, the NYPD has placed one detective on suspension and one supervisor on modified assignment status pending further investigation into the matter...</p>
<b>Summary</b>	<p>Store manager claims Detective Ian Cyrus was caught on camera stealing . New York detectives were arresting workers for selling untaxed cigarettes . He's been suspended without pay and his supervisor placed on desk duty .</p>
Icelandic Translation	
<b>Document</b>	<p>A New York einkaspæjara hefur verið frestað eftir að hafa verið sakaður um að ræna verslun á meðan samstarfsmenn voru handtaka starfsmenn fyrir að selja ólöglega sígarettur. Framkvæmdastjóri sælkeraverslunarinnar í Brooklyn segir að rannsóknarlögreglumadurinn Ian Cyrus sem hefur verið leystur frá störfum án launa hafi verið myndaður við að stela peningum. Yfirmaður hans hefur verið settur á skrifborðsvakt. Skrolla niður fyrir vídeó. Náðist á mynd: Við skoðun á öryggismyndavélum verslunarinnar sagðist verslunarstjórinn hafa séð einn rannsóknarlögreglumannanna finna peningakassann undir afgreiðsluborðinu og rannsóknarlögreglumanninn Ian Cyrus grípa handfylli af honum. Að sögn Abdullah beygði rannsóknarlögreglumadurinn Ian Cyrus sig niður og þegar hann stóð upp virtist hann setja peningana í vasann á jakkanum sínum. Fimm rannsóknarlögreglumenn frá lögreglunni í New York fóru í verslunina í Brooklyn föstudaginn 3. apríl og handtóku tvo starfsmenn sem voru sakaðir um að selja óskattlagðar sígarettur samkvæmt frétt ABC7 Eyewitness News. Lögreglumennirnir eru sagðir hafa tekið peninga úr afgreiðslukassanum og sígarettupakka sem sönnunargögn. Framkvæmdastjórinn Ali Abdullah sem var ekki í búðinni á þeim tíma fékk kvittun frá lögreglunni daginn eftir þar sem stóð að 593 dollarar hefðu verið teknir en hann áttaði sig á því að 2.650 dollara vantaði upp á leiguna. Við skoðun á öryggismyndavélum sagði hann að Cyrus lögreglumadur hefði tekið handfylli af peningum úr peningakassa undir afgreiðsluborðinu. Að sögn Abdullah beygði lögreglumadurinn sig niður og þegar hann stóð upp virtist hann setja peningana í vasann á jakkanum sínum. "Ég lít á kerfið mitt og sé að lögreglumadurinn tók peningana" sagði Abdullah í samtali við fréttastofu. "Þetta er brjálæði" Hann segist hafa hringt í nærliggjandi lögreglustöð sem sendi eftirlitsmann til að skoða myndbandið. Abdullah heldur því fram að yfirmaðurinn hafi sagt: "Hann mun missa lífeyrinn sinn, missa lífið, fyrir 2.600 dollara?" Stephen Davis, aðstoðarlögreglustjóri hjá lögreglunni í New York, sagði í samtali við MailOnline: "Miðað við eðli ásakana í þessu atviki, auk myndbandsins sem okkur var afhent, hefur lögreglan í New York sett einn rannsóknarlögreglumann í tímabundið leyfi og einn yfirmann í breytt verkefni á meðan frekari rannsókn fer fram á málinu"...</p>
<b>Summary</b>	<p>Verslun framkvæmdastjóri krafa Leynilögreglumadur Ian Cyrus var veiddur á myndavél stela. New York lögreglumönnum voru handtaka starfsmenn til að selja óskattlagðar sígarettur. Hann hefur verið frestað án launa og yfirmaður hans settur á skrifborðið vakt.</p>

## B. Hyperparameters

Pre-training of Gullfaxi in Figure 4.1				
Model	# Steps	Batch size	Max input tokens	Max target tokens
Gullfaxi <sub>BASE</sub>	200k	256	512	256
Gullfaxi <sub>LARGE</sub>	135k	8192	512	256

Fine-tuning of Gullfaxi models in Table 4.1, 4.2, 4.3 and 4.4				
Task	# Steps	Batch size	Max input tokens	Max target tokens
Intro + Main → Summary	4k	256	512	128
Main → Intro	4k	256	512	128
Intro → Title	4k	256	128	32
XSum	50k	256	512	64
CNN/DailyMail	50k	256	512	128

Fine-tuning of mT5 <sub>BASE</sub> in Table 4.1 and 4.4				
Task	# Steps	Batch size	Max input tokens	Max target tokens
RRN	8k	256	Same as Gullfaxi	

Low-resource fine-tuning of Gullfaxi <sub>BASE</sub> and mT5 <sub>BASE</sub> in Figure 4.2				
Task	# Steps	Batch size	Max input tokens	Max target tokens
RRN	3k	256	Same as Gullfaxi	



## C. Example Model Outputs

Example outputs from all models for a corresponding RRN test set document as well as their ROUGE-F1 (R1/R2/RL) scores. The outputs presented are from models that have been fine-tuned on a corresponding RRN task.

## C. Example Model Outputs

RRN Document		
<b>Title</b>	Neyðarástand vegna flóða í Kína	
<b>Intro</b>	Hæsta viðbúnaðarstigi hefur verið lýst yfir í Henan-héraði í Kína vegna flóða. Þau hafa orðið að minnsta kosti tólf manns að bana.	
<b>Main</b>	Hátt í tvö hundruð þúsund íbúar borgarinnar Sheng-sjá Zhengzhou í Henan-héraði í Kína hafa verið fluttir að heiman vegna flóða. Þau hafa orðið að minnsta kosti tólf manns að bana. Hæsta viðbúnaðarstigi hefur verið lýst yfir í héraðinu. Ríkisfjölmiðlar í Kína hafa eftir Xi Jinping að ástandið í Henan sé afar alvarlegt. Stíflur hafi brostið og valdið manntjóni og eignatapi. Allir verði að leggjast á árarar til að koma í veg fyrir að það verði enn meira. Á annan tug borga og bæja eru umflotin vatni. Á götum hafa myndast straumharðar ár sem bera með sér bíla og alls kyns brak. Ef marka má fréttir af svæðinu er ástandið verst í héraðshöfuðborginni Zhengzhou. Þar flæddi vatn inn í jarðlestagöng með þeim afleiðingum að tólf drukkuðu. Um fimm hundruð var bjargað úr göngunum. Hátt í tvö hundruð þúsund íbúum borgarinnar hefur verið forðað að heiman vegna flóða. Síðustu þrjá sólarhringa hefur fallið álíka mikið regn og á einu ári. Þá greindi kínverski herinn frá því í gær að tuttugu metra sprunga væri komin í Yihetan stífluna í Luoyang þar sem um það bil sjö milljónir búa. Hermenn hafa verið sendir á vettvang til að stýra rennsli í ám og hlaða upp sandpokum til að styrkja bakkana. Vegna veðursins í miðhluta Kína hefur fjöldi flugferða verið aflýst og áætlanarferðir járnbrautarlesta eru úr skorðum.	
<b>Summary</b>	Hæsta viðbúnaðarstigi hefur verið lýst yfir í Henan héraði í Kína vegna flóða. Að minnsta kosti tólf manns hafa látist.	
Gullfaxi <sub>BASE</sub> (RRN)		ROUGE F1
<b>Title</b>	skjálfti í henan-héraði	25.00/0.00/25.00
<b>Intro</b>	að minnsta kosti fimm eru látnir eftir að öflugur jarðskjálfti reið yfir kína í gær.	35.29/11.43/29.41
<b>Summary</b>	minnst níu hafa látist í miklum flóðum í kína frá því í gær. tugir þúsunda hafa orðið að yfirgefa heimili sín vegna flóða.	35.00/14.29/30.00
Gullfaxi <sub>BASE</sub> (CNN/DailyMail)		ROUGE F1
<b>Intro</b>	að minnsta kosti 12 manns hafa verið drepnir í henan héraði í henan-héraði í kína	54.55/28.57/48.48
<b>Summary</b>	að minnsta kosti 12 manns hafa látist í henan héraði í henan-héraði í henan-héraði í kína	64.52/50.00/64.52
Gullfaxi <sub>BASE</sub> (CNN/DailyMail + RRN)		ROUGE F1
<b>Intro</b>	hátt í tvö þúsund íbúar í kína hafa orðið að yfirgefa heimili sín vegna flóða í henan-héraði í kína.	45.71/31.58/40.00
<b>Summary</b>	hæsta viðbúnaðarstigi hefur verið lýst yfir í kína vegna flóða í henan-héraði í kína vegna flóða. þau hafa orðið að minnsta kosti tólf að bana í henan-héraði í kína.	78.95/58.54/73.68
Gullfaxi <sub>LARGE</sub> (RRN)		ROUGE F1
<b>Title</b>	flóð í christchurch	25.00/0.00/25.00
<b>Intro</b>	að minnsta kosti einn lét lífið í flóðum og aurskriðum í kína í gær. þúsundir hafa þurft að yfirgefa heimili sín.	31.58/14.63/31.58
<b>Summary</b>	neyðarástandi hefur verið lýst yfir á flóðasvæðunum í henan héraði í henan héraði í suðurhluta landsins.	43.75/37.50/43.75.
mT5 <sub>BASE</sub> (RRN)		ROUGE F1
<b>Title</b>	Hæsta viðbúnaðarstig lýst yfir í Henan	18.18/0.00/18.18
<b>Intro</b>	Síðustu þrjá sólarhringa hefur fallið ástandið verst í miðhluta Kína. Hátt í tvö hundruð þúsund íbúar borgarinnar hafa verið fluttir að heiman vegna flóða.	37.21/4.55/32.56
<b>Summary</b>	Hæsta viðbúnaðarstigi hefur verið lýst yfir í Henan-héraði í Kína vegna flóða. Þau hafa orðið að minnsta kosti tólf manns að bana.	82.05/65.00/76.92

RRN Document

<b>Title</b>	Stríðið gæti varað árum saman	
<b>Intro</b>	Stríðið í Úkraínu gæti varað árum saman og Vesturlönd þurfa að vera undir það búin að veita Úkraínu stuðning áfram. Þetta segir Jens Stoltenberg framkvæmdastjóri Atlantshafsbandalagsins.	
<b>Main</b>	Stoltenberg lét þessi orð falla í viðtalið við þýska dagblaðið Bild. Hann segir að þrátt fyrir að stríðið sé kostnaðarsamt verði Vesturlönd að halda áfram að senda Úkraínumönnum hergögn og nauðsynjar því kostnaðurinn verði umtalsvert meiri ef Vladímír Pútín Rússlandsforseti nái sínum markmiðum. Við þurfum að búa okkur undar það að stríðið gæti varað árum saman, segir Stoltenberg. Það sama segir Boris Johnson, forsætisráðherra Bretlands, sem fór í sína aðra heimsókn til Kænugarðs á föstudag. "It would be a catastrophe if Putin won. It would be a catastrophe if he was able to secure the land bridge to the cities in the south that he has, to hold the Donbas. That's what he wants. Boris Johnson segir skelfilegt að hugsa til þess að Pútín vinni stríðið. Stuðningur við Úkraínu sé nauðsynlegur til að koma í veg fyrir að Rússar nái Donbas. Breska varnarmálaráðuneytið birti í morgun yfirferð um átökinn sem geisa hvað harðast við borgina Sjevjerodonetsk. Þar kemur fram að litlar breytingar hafi orðið síðasta sólarhringinn. Síðustu daga hafi verið nokkuð um liðhlaup úr úkraínskum hersveitum. Einnig segir breska varnarmálaráðuneytið að átök haldi áfram innan hersveita Rússa og dæmi séu um að hersveitir neiti að hlýða skipunum foringja sinna.	
<b>Summary</b>	Framkvæmdastjóri Atlantshafsbandalagsins segir að Vesturlönd þurfi að vera undir það búin að veita Úkraínu stuðning áfram, stríðið geti varað í mörg ár í viðbót.	
<b>Gullfaxi<sub>BASE</sub> (RRN)</b>		<b>ROUGE F1</b>
<b>Title</b>	jens stoltenberg um úkraínu	0.00/0.00/0.00
<b>Intro</b>	vladímír pútín rússlandsforseti segir að stríðið í úkraínu sé í fullum gangi. pútín rússlandsforseti segir að rússar verði að sætta sig við stríðið í úkraínu.	25.00/9.09/25.00
<b>Summary</b>	framkvæmdastjóri atlantshafsbandalagsins segir að stríðið í úkraínu sé hafið. framkvæmdastjóri atlantshafsbandalagsins segir að stríðið í úkraínu sé hafið.	40.00/6.25/33.33
<b>Gullfaxi<sub>BASE</sub> (CNN/DailyMail)</b>		<b>ROUGE F1</b>
<b>Intro</b>	forsætisráðherra bretlands segir að stríðið í úkraínu sé nauðsynlegurnauðsynlegur til að koma í veg fyrir að rússland nái árangri í úkraínu	25.00/9.09/25.00
<b>Summary</b>	jens stoltenberg, framkvæmdastjóri atlantshafsbandalagsins í úkraínu, segir að vladímírs pútín rússlandsforseti muni halda áfram að senda hergögn til kænugarðs. forsætisráðherra bretlands, donbas, segir að rússneskir hermenn hafi verið sendir til kænugarðs til að koma í veg fyrir átök í úkraínu.	19.23/3.33/15.38
<b>Gullfaxi<sub>BASE</sub> (CNN/DailyMail + RRN)</b>		<b>ROUGE F1</b>
<b>Intro</b>	forsætisráðherra bretlands segir skelfilegt að stríðið í úkraínu sé að undirbúa sig fyrir það. hann segir að stríð í úkraínu sé á réttri leið.	28.57/8.70/28.57
<b>Summary</b>	jens stoltenberg knattspyrnustjóri atlantshafsbandalagsins segir að stríðið í úkraínu gæti komið í veg fyrir að rússar nái friði í úkraínu.	26.32/4.76/21.05
<b>Gullfaxi<sub>LARGE</sub> (RRN)</b>		<b>ROUGE F1</b>
<b>Title</b>	staðan í úkraínu	0.00/0.00/0.00
<b>Intro</b>	boris jeltsín rússlandsforseti segir að rússar þurfi að grípa til aðgerða til að koma í veg fyrir að rússar dragi herlið sitt til baka.	14.29/0.00/9.52
<b>Summary</b>	jens stoltenberg, forsætisráðherra bretlands, segir að stríðið í úkraínu sé í hættu. hann segir að rússar verði að sætta sig við stríðið í úkraínu.	26.32/4.65/21.05.
<b>mT5<sub>BASE</sub> (RRN)</b>		<b>ROUGE F1</b>
<b>Title</b>	Stríðið í Úkraínu	25.00/0.00/25.00
<b>Intro</b>	Forsætisráðherra Bretlands segir skelfilegt að hugsa til þess að Vladímír Pútín Rússlandsforseti nái sínum markmiðum.	10.53/0.00/10.53
<b>Summary</b>	Forsætisráðherra Bretlands segir að stríðið í Úkraínu gæti varað árum saman og Vesturlönd þurfa að vera undir það búin að veita Úkraínu stuðning áfram. Þetta segir framkvæmdastjóri Atlantshafsbandalagsins.	62.22/36.00/44.44



### C. Example Model Outputs

RRN Document		
<b>Title</b>	Breytingar á leigumarkaði	
<b>Intro</b>	Ungt fólk hefur hrakist af leigumarkaði í covid-faraldrinum og hefur í vaxandi mæli þurft að flytja aftur heim í foreldrahús. Vísbendingar eru um að dregið hafi úr framboði á leiguhúsnæði á síðustu mánuðum.	
<b>Main</b>	Þetta kemur fram í könnun á vegum hagdeildar Húsnæðis- og mannvirkjastofnunar um stöðuna á leigumarkaði. Almennt hefur leigjendum fækkað á síðustu tveimur árum. Margir hafa nýtt sér hagstæð lánakjör til að kaupa húsnæði en aðrir hafa hrakist af markaðinum. Karlotta Halldórsdóttir hagfræðingur hjá stofnuninni segir að ungt fólk hafi orðið hvað verst úti. Svo myndi ég segja líka að áhrif af covid faraldrinum þau eru að koma einna verst niður á búsetu hjá ungu fólki. Yngsta aldurshópnun 18 til 24 ára. Þar erum við að sjá stórlækkun hjá þeim á leigumarkaði og þeir eru í auknum mæli að flytja aftur í foreldrahús. Við mundum geta giskað á að það væri minni atvinna fyrir þau. Kannski mikið til fólk sem er í hlutastörfum, veitingageiranum, ferðaþjónustutengt og atvinna þar hafi minnkað og þau séu með minna ráðstöfunarfé og hafi þá ekki lengur efni á að vera á leigumarkaði. Þá eru einnig vísbendingar um að dregið hafi úr framboði á leiguhúsnæði á undanförnum mánuðum. Í upphafi faraldursins jókst framboð verulega þegar Airbnb íbúðir fóru inn á hinn almenna leigumarkað. Við sáum einmitt í kjölfarið af covid og í fyrra þá vorum við að sjá framboð af leiguhúsnæði aukast töluvert. Fólk átti auðveldara með að finna sér húsnæði og svona auknið öryggi. En við erum að sjá í fyrsta skipti núna síðan 2015 að það sé hærra hlutfall sem finnst erfiðara að verða sér úti um húsnæði. Þannig að það eru vísbending um að það sé að draga úr þessu mikla framboði sem var Þetta hefur líka þær afleiðingar að leiguverð fer nú hækkandi. Við sáum það í fyrsta skipti í svoltinn tíma þá hækkaði vísitalan. alveg töluvert milli mánaða. Þannig að það eru kannski svona einhver merki um að það sé aðeins farið að vera erfiðari leigumarkaðurinn hjá okkur.	
<b>Summary</b>	Staða ungs fólks á húsnæðismarkaði hefur versnað í Covid faraldrinum og margir hafa neyðst til að flytja aftur í foreldrahús. Þetta sýnir ný könnun.	
Gullfaxi <sub>BASE</sub> (RRN)		ROUGE F1
<b>Title</b>	skortur á leiguhúsnæði	33.33/0.00/33.33
<b>Intro</b>	ungtungt fólk á aldrinum 18 til átján ára hefur flutt lögheimili sitt á almennum leigumarkaði. Þetta kemur fram í nýrri könnun.	20.83/0.00/20.83
<b>Summary</b>	ungt fólk hefur dregið úr framboði á leigumarkaði frá því í fyrra. ungt fólk er líklegra til að flytjast búferlum frá heimilum sínum.	26.67/4.35/22.22
Gullfaxi <sub>BASE</sub> (CNN/DailyMail)		ROUGE F1
<b>Intro</b>	ungt fólk á aldrinum 18 til 24 ára hefur fækkað á undanförunum tveimur árum. ungt fólk á aldrinum 18 til 24 ára eru verst úti á leigumarkaði.	26.67/3.92/22.22
<b>Summary</b>	ungt fólk hefur dregið úr framboði á undanförunum mánuðum. á undanförunum mánuðum hefur fjöldi leigjenda dregið úr framboði á leiguhúsnæði á undanförunum mánuðum.	10.81/0.00/10.81
Gullfaxi <sub>BASE</sub> (CNN/DailyMail + RRN)		ROUGE F1
<b>Intro</b>	ungt fólk á aldrinum 18 til 24 ára er í auknum mæli á leigumarkaði í fyrsta sinn í 20 ár. Þetta kemur fram í nýrri könnun sem gerð var á vegum húsnæðis- og mannvirkjastofnunar sem unnin var á dögunum.	23.73/5.80/20.34
<b>Summary</b>	ungt fólk hefur hrakist af leigumarkaði í vaxandi mæli á síðustu mánuðum. vísbendingar eru um að ungt fólk hafi dregið úr framboði covid-faraldrinum.	21.74/0.00/17.39
Gullfaxi <sub>LARGE</sub> (RRN)		ROUGE F1
<b>Title</b>	ungt fólk á leigumarkaði	57.14/40.00/57.14
<b>Intro</b>	framkvæmdastjóri félags fasteignasala segir að skortur sé á húsnæði á viðráðanlegu verði fyrir ungt fólk. framkvæmdastjóri félags fasteignasala segir að skortur sé á leiguhúsnæði.	23.26/8.33/23.26
<b>Summary</b>	ungt fólk er líklegra til að vera á leigumarkaði á viðráðanlegu verði. ungt fólk er líklegra til að vera á leigumarkaði.	22.22/5.41/16.67.
mT5 <sub>BASE</sub> (RRN)		ROUGE F1
<b>Title</b>	Leigumarkaður	0.00/0.00/0.00
<b>Intro</b>	Forstjóri Húsnæðis- og mannvirkjastofnunar segir að ungt fólk hafi orðið hvað verst úti um húsnæði. Þetta kemur fram í könnun á vegum hagdeildar Húsnæðis- og mannvirkjastofnunar um stöðuna á leigumarkaði.	33.96/3.39/18.87
<b>Summary</b>	Formaður Húsnæðis- og mannvirkjastofnunar segir að ungt fólk hefur hrakist af leigumarkaði í covid-faraldrinum og hefur í vaxandi mæli þurft að flytja aftur heim í foreldrahús. Vísbendingar eru um að dregið hafi úr framboði á leiguhúsnæði á síðustu mánuðum.	28.57/9.68/25.00

RRN Document		
<b>Title</b>	Jens vill verða seðlabankastjóri	
<b>Intro</b>	Jens Stoltenberg, framkvæmdastjóri Atlantshafsbandalagsins, er meðal tuttugu og tveggja umsækjenda um starf seðlabankastjóra Noregs.	
<b>Main</b>	Tuttugu og tvær umsóknir bárust um embætti seðlabankastjóra Noregs. Eitt nafn vekur meiri athygli en önnur. Jens Stoltenberg, framkvæmdastjóri Atlantshafsbandalagsins og fyrrverandi forsætisráðherra Noregs, er meðal umsækjenda. Í tilkynningu frá blaðafulltrúa Stoltenbergs segir að norska fjármálaráðuneytið hafi haft samband við hann í síðasta mánuði og spurt hvort hann gæti hugsað sér að sækja um starfið. Hann hafi svarað því játandi, en ef svo færi að hann yrði fyrir valinu gæti hann ekki hafið störf fyrr en fyrsta október, þegar veru hans hjá Atlantshafsbandalaginu lýkur. Øystein Olsen hefur gegnt gegnt embætti seðlabankastjóra frá því í janúar 2011. Hann hefur óskað eftir að láta af störfum fyrsta febrúar næstkomandi, þegar hann nær eftirlaunaaldri. Af þeim tuttugu og tveimur sem sækja um stöðu seðlabankastjóra er ein kona, Ida Wolden Bache aðstoðar-seðlabankastjóri. Á norska viðskiptafréttavefnum E24 er haft eftir henni að hún hafi sótt um eftir að hafa ráðfært sig við fjármálaráðuneytið, sem ræður í starfið. Hún viti vel hvað felist í því að vera seðlabankastjóri og geti vel hugsað sér að takast á við það.	
<b>Summary</b>	Jens Stoltenberg, framkvæmdastjóri Atlantshafsbandalagsins, er meðal tuttugu og tveggja umsækjenda um starf seðlabankastjóra Noregs.	
Gullfaxi <sub>BASE</sub> (RRN)		ROUGE F1
<b>Title</b>	stoltenberg í nato	0.00/0.00/0.00
<b>Intro</b>	jens stoltenberg, forsætisráðherra noregs, tilkynnti í morgun að hann hygðist bjóða sig fram í embætti seðlabankastjóra.	20.69/7.14/20.69
<b>Summary</b>	jens stoltenberg, framkvæmdastjóri atlantshafsbandalagsins (nato), hefur sagt af sér sem framkvæmdastjóri atlantshafsbandalagsins. hann hefur sagt af sér sem framkvæmdastjóri atlantshafsbandalagsins.	29.63/14.81/29.63
Gullfaxi <sub>BASE</sub> (CNN/DailyMail)		ROUGE F1
<b>Intro</b>	jens stoltenberg, fyrrum forsætisráðherra noregs, hefur óskað eftir því að norska fjármálaráðuneytið rannsaki hann. norska fjármálaráðuneytið heldur því fram að norska fjármálaráðuneytið hafi haft samband við hann í síðasta mánuði.	10.81/5.13/10.81
<b>Summary</b>	jens stoltenberg, framkvæmdastjóri atlantshafsbandalagsins og fyrrverandi forsætisráðherra noregs, hefur verið ráðinn í starfið. norska fjármálaráðuneytið hefur haft samband við hann í síðasta mánuði.	22.22/11.11/22.22
Gullfaxi <sub>BASE</sub> (CNN/DailyMail + RRN)		ROUGE F1
<b>Intro</b>	jens stoltenberg, framkvæmdastjóri atlantshafsbandalagsins og fyrrverandi forsætisráðherra, hefur sagt af sér eftir að umsóknir bárust um að sækja um stöðuna í febrúar.	29.41/11.76/29.41
<b>Summary</b>	jens stoltenberg, framkvæmdastjóri atlantshafsbandalagsins og fyrrverandi forsætisráðherra, er meðal tuttugu umsækjenda um starfið. jens stoltenberg, framkvæmdastjóri atlantshafsbandalagsins er meðal umsækjenda um starf.	73.33/38.71/66.67
Gullfaxi <sub>LARGE</sub> (RRN)		ROUGE F1
<b>Title</b>	jens stoltenberg lætur af störfum	22.22/0.00/22.22
<b>Intro</b>	jens stoltenberg, forsætisráðherra noregs, hefur skipað jens stoltenberg, forsætisráðherra noregs.	28.57/10.00/28.57
<b>Summary</b>	jens stoltenberg, fyrrverandi forsætisráðherra noregs, hefur verið ráðinn nýr seðlabankastjóri.	23.08/8.33/23.08.
mT5 <sub>BASE</sub> (RRN)		ROUGE F1
<b>Title</b>	Nýr leiðtogi í NATO	0.00/0.00/0.00
<b>Intro</b>	Tveir umsóknir bárust um embætti seðlabankastjóra Noregs. Eitt nafn vekur meiri athygli en önnur.	21.43/7.69/21.43
<b>Summary</b>	Prverandi forsætisráðherra Noregs er meðal tuttugu og tveggja umsækjenda um starf seðlabankastjóra Noregs. Eitt nafn vekur meiri athygli en önnur.	58.82/54.55/58.82

### C. Example Model Outputs

RRN Document		
<b>Title</b>	Veður hér verður vott í framtíðinni	
<b>Intro</b>	Úrkoma verður meiri hér á landi í framtíðinni og þá frekar í formi rigningu en snjókomu. Þetta eru niðurstöður ítarlegra rannsókna Landsvirkjunar á horfum í vatnabúskap eftir að jöklar hafa bráðnað.	
<b>Main</b>	Veður á Íslandi verður vott í framtíðinni, svipað og á Skotlandi. Þetta leiða rannsóknir vísindamanna Landsvirkjunar í ljós. Hörður Arnarson forstjóri Landsvirkjunar segir að bráðnun jökla hafi áhrif á vatnabúskap og vatnsaflsvirkjanir. Við höfum lagt mikinn metnað í það á síðustu áratugum, alveg frá því fyrir aldamót, sem vísindamenn Landsvirkjunar fóru í umfangsmiklar rannsóknir á því hvers megi vænta. Við tókum þátt í alþjóðlegum rannsóknarverkefnum og höfum birt mikið af ritrýndum rannsóknum um áhrifin. Þannig að við teljum okkur þekkja býsna vel áhrifin. Til skemmri tíma mun þetta hafa þau áhrif, meðan jöklarnir eru að bráðna því miður, við gerum ráð fyrir að það sé bara staðreynd sem við erum að horfa á, þá mun það þýða aukna orkuvinnslu því við munum sjá það gerast miðað við þann hraða sem við erum að spá. Síðan tekur við tímabil þar sem við förum aftur í ástandið sem við vorum með fyrir breytingar, með svipaða orkugetu. En þá mun veðurfarid líklega hafa breyst. Við verum þá líklega með votviðrasamara veðurfar er talið, ekki ósvipað og við sjáum á Skotlandi en jafnvel meiri úrkoma. En hún mun falla meira sem rigning og minna sem snjór. Þannig að það mun verða á suman hátt auðveldara að reka kerfið. - En kannski leiðinlegra að búa hérna? Já, auðvitað eru þetta mjög slæmar breytingar. Þó að við séum að horfast í augu við þær í þessum rannsóknum þá eigum við að gera allt til að koma í veg fyrir þær. En þessar rannsóknir miða fyrst og fremst við þetta ef þetta gerist. En þetta er náttúrulega ekki það sem við viljum og við þurfum að aðlaga okkur að þeim. Og við sem orkufyrirtæki Íslendinga leggjum mikla áherslu á að aðlaga okkur að þessum breytingum og að þær komi okkur ekki á óvart. Þessar rannsóknir okkar vísindamanna hafa vakið mikla athygli erlendis og ég held að það séu fá vatnsorkufyrirtæki í heiminum sem hafa lagt í jafnmiklar rannsóknir og þekkja áhrifin jafnvel og við gerum. Þannig að við erum í þessu eins og öðru að reyna að vanda okkur, gera vel og vinna út frá þeim forsendum sem eru í umhverfinu.	
<b>Summary</b>	Úrkoma verður meiri hér á landi í framtíðinni og þá frekar í formi rigningar en snjókomu. Þetta eru niðurstöður ítarlegra rannsókna Landsvirkjunar á horfum í vatnabúskap eftir að jöklar bráðna.	
Gullfaxi <sub>BASE</sub> (RRN)		ROUGE F1
<b>Title</b>	úrkoma í jökulsá á fjöllum	18.18/0.00/18.18
<b>Intro</b>	landsvirkjun hefur ákveðið að grípa til aðgerða til að sporna við loftslagsbreytingum og grípa til aðgerða til að draga úr líkum á vatnavöxtum.	13.64/0.00/9.09
<b>Summary</b>	prófessor í jarðeðlisfræði segir að þetta sé í fyrsta skipti sem mælingar á jöklum á Íslandi hafa borið á góma. Þetta segir forstjóri landsvirkjunar.	25.00/0.00/20.83
Gullfaxi <sub>BASE</sub> (CNN/DailyMail)		ROUGE F1
<b>Intro</b>	bráðnun jökla í skotlandi mun leiða í ljós hvernig á að takast á við loftslagsbreytingar.	14.63/0.00/9.76
<b>Summary</b>	bráðnun ís hefur áhrif á vatnabúskap í skotlandi. bráðnun ís í skotlandi mun hafa áhrif á bráðnun ís í skotlandi. bráðnun ís hefur áhrif á bráðnun ís í skotlandi.	15.79/0.00/15.79
Gullfaxi <sub>BASE</sub> (CNN/DailyMail + RRN)		ROUGE F1
<b>Intro</b>	nýlegar rannsóknir hafa sýnt fram á að jöklaframleiðsla minnki á svipaðan hátt og talið er að hún hafi kostað á svipaðan hátt og í skotlandi.	21.28/0.00/17.02
<b>Summary</b>	úrkoma verður meiri hér á landi en í framtíðinni og í framtíðinni. Þetta eru niðurstöðurnar í nýrri rannsókn á því hvers vegna jöklar hafa bráðnað.	53.06/30.19/48.98
Gullfaxi <sub>LARGE</sub> (RRN)		ROUGE F1
<b>Title</b>	úrkoma í kortunum	22.22/0.00/22.22
<b>Intro</b>	nýjar rannsóknir á jöklum og jöklum á Íslandi hafa leitt í ljós að íshellan á jöklum hefur ekki áhrif á loftslag. Þetta segir forstjóri landsvirkjunar.	29.17/0.00/25.00
<b>Summary</b>	veðurstofan spáir hlýindum á næstu dögum.	11.43/0.00/11.43.
mT5 <sub>BASE</sub> (RRN)		ROUGE F1
<b>Title</b>	Landsvirkjun um snjókomu	0.00/0.00/0.00
<b>Intro</b>	Forstjóri Landsvirkjunar segir að bráðnun jökla hafi áhrif á vatnsaflsvirkjanir.	15.79/0.00/15.79
<b>Summary</b>	Úrkoma verður meiri hér á landi í framtíðinni og þá frekar í formi rigningu en snjókomu. Þetta eru niðurstöður ítarlegra rannsókna Landsvirkjunar á horfum í vatnabúskap eftir að jöklar hafa bráðnað.	91.23/85.25/91.23

**RRN Document**

<b>Title</b>	Stefan Löfven segir af sér
<b>Intro</b>	Stefan Löfven, forsætisráðherra Svíþjóðar, ætlar að láta af embætti. Sænska þingið lýsti yfir vantrausti á stjórn hans í síðustu viku.
<b>Main</b>	Stefan Löfven, forsætisráðherra Svíþjóðar, hefur ákveðið að láta af embætti forsætisráðherra. Vantrausti var lýst á stjórn hans í síðustu viku. Forseti sænska þingsins kallar leiðtoga allra flokka á þingi á sinn fund áður en næstu skref verða ákveðin. Löfven tilkynnti ákvörðun sína á fundi með fréttamönnum í morgun. Hann ákvað jafnframt að boða ekki til aukakosninga en sænsk lög kveða á um að kosið sé á fjögurra ára fresti, hvað sem á gengur. Kjósa á haustið 2022. Löfven sagði að aukakosningar myndu hugsanlega draga tímabil óvissu á langinn og við það yrði ekki unað á tímum heimsfaraldurs. Alltaf yrði að gera það sem væri Svíþjóð fyrir bestu. Löfven sagðist hafa komist að þeirri niðurstöðu að segja af sér um helgina. Það væri erfiðasta pólitíska ákvörðun sem hann hefði þurft að taka en hann teldi þetta fyrir bestu. Ákvörðun hans um afsögn en að boða ekki til aukakosninga var nokkuð óvænt. Um fjóra mánuði tók að mynda stjórn eftir kosningarnar 2018. Andreas Norlén þingforseti kallar á næstunni leiðtoga allra flokka sem sæti eiga á þingi á sinn fund til að kanna möguleikana á að mynda nýja stjórn. Takist það ekki eftir fjórar tilraunir verður boðað til aukakosninga.
<b>Summary</b>	Stefan Löfven, forsætisráðherra Svíþjóðar, ætlar að láta af embætti. Sænska þingið lýsti yfir vantrausti á stjórn hans í síðustu viku.

	Gullfaxi <sub>BASE</sub> (RRN)	ROUGE F1
<b>Title</b>	löfven, forsætisráðherra svíþjóðar	0.00/0.00/0.00
<b>Intro</b>	stefan löfven, forsætisráðherra svíþjóðar, tilkynnti í morgun að hann ætlaði að bjóða sig fram til endurkjörs í næstu kosningum. hann tilkynnti í morgun að hann ætlaði ekki að bjóða sig fram til endurkjörs.	31.58/14.63/31.58
<b>Summary</b>	stefan löfven forsætisráðherra ætlar að láta af embætti forsætisráðherra svíþjóðar í næstu viku.	54.55/25.00/54.55

	Gullfaxi <sub>BASE</sub> (CNN/DailyMail)	ROUGE F1
<b>Intro</b>	sænska pm tilkynnti ákvörðun sína á fundi í morgun. pm tilkynnti ákvörðun sína á fundi með fréttamönnum í morgun.	19.35/0.00/19.35
<b>Summary</b>	stefan ven, forsætisráðherra svíþjóðar, hefur lýst yfir vantrausti á forsætisráðherra svíþjóðar. forsætisráðherrann lýsti yfir vantrausti á forsætisráðherrann í síðustu viku.	57.14/33.33/57.14

	Gullfaxi <sub>BASE</sub> (CNN/DailyMail + RRN)	ROUGE F1
<b>Intro</b>	stefan löfven forsætisráðherra svíþjóðar hefur ákveðið að segja af sér embætti forsætisráðherra svíþjóðar. hann hefur ákveðið að segja af sér embætti forsætisráðherrans á fjögurra ára fresti.	32.43/0.00/32.43
<b>Summary</b>	stefan löfven, forsætisráðherra svíþjóðar, ætlar að láta af störfum á þingi í kjölfar vantrausti á sænska þingið í síðustu viku.	76.92/56.41/66.67

	Gullfaxi <sub>LARGE</sub> (RRN)	ROUGE F1
<b>Title</b>	löfven, forsætisráðherra svíþjóðar	0.00/0.00/0.00
<b>Intro</b>	stefan löfven, forsætisráðherra svíþjóðar, tilkynnti í dag að hann ætli að bjóða sig fram til formennsku í flokknum. löfven tilkynnti í morgun að hann hygðist bjóða sig fram til formennsku í flokknum.	30.77/14.29/30.77
<b>Summary</b>	stefan löfven, forsætisráðherra svíþjóðar, tilkynnti í morgun að stefan löfven, forsætisráðherra svíþjóðar, verði næsti forsætisráðherra svíþjóðar.	37.50/18.75/37.50.

	mT5 <sub>BASE</sub> (RRN)	ROUGE F1
<b>Title</b>	Löfven fær ekki af sér	60.00/25.00/60.00
<b>Intro</b>	Forsætisráðherra Svíþjóðar hefur ákveðið að láta af sér embætti forsætisráðherra. Vantrausti var lýst á stjórn hans í síðustu viku.	63.16/37.84/63.16
<b>Summary</b>	Löfven, forsætisráðherra Svíþjóðar, ætlar að láta af embætti. Sænska þingið lýsti yfir vantrausti á stjórn hans í síðustu viku.	97.44/97.30/97.44

### C. Example Model Outputs

RRN Document		
<b>Title</b>	Eiríkur Bergmann	
<b>Intro</b>	Stjórn málafræðingur segir að sú gagnrýni sem komið hefur fram vegna sölunnar á hlut ríkisins í Íslandsbanka muni reyna enn frekar á stjórnarsamstarfið. Ríkisstjórnin sé nú stödd í miðjum hvirfilbyl og að lítið þurfi til að slitni upp úr.	
<b>Main</b>	Ríkisstjórnin hefur sætt harðri gagnrýni fyrir söluna á Íslandsbanka og mótmæli verið haldin nær vikulega, síðast í gær þegar fjöldi fólks kom saman á Austurvelli og krafðist þess að sölu yrði rift og að ríkisstjórnin færi frá völdum. Eiríkur Bergmann stjórn málafræðingur segir að flækjustigið við að rifta sölnni sé töluvert en að augljóst sé að titrings gæti á stjórnarheimilinu. Ríkisstjórnin hefur nokkuð skyndilega lent í einhvers konar hvirfilbyl sem skukur hana all nokkuð að því er virðist. Manni virðist sem það sé orðið nokkuð útbreidd skoðun í íslensku þjóðfélagi að það hafi verið verulegar brotalamir í þessu útboði Íslandsbanka og það mál er einfaldlega þannig að það getur hæglega klofið pólitíska afstöðu milli stjórnarflokkanna Staðan sé einna erfiðust fyrir bakland Vinstri grænna enda hefði sá flokkur undir öðrum kringumstæðum gagnrýnt söluna harðlega. Þegar mótmæli gerast svona útbreidd eins og við höfum orðið vitni að að þá verður þetta mál pólitískt erfiðast fyrir vinstri græna og af þeim sökum nötr ríkisstjórnarborðið. Óvíst sé hversu mikið meira ríkisstjórnin þoli. Í ruaninni er þetta í fyrsta sinn sem við sjáum einhvers konar brotsjó ganga yfir stjórn nina. Og hvort að fleyið sigli í gegnum ólgusjóinn vitum við auðvitað ekki enn þá en mér synkir þetta edkki endilea þannig gjörningaveður að það sökkti fleyinu þegar upp er staðið.	
<b>Summary</b>	Stjórn málafræðingur segir að ríkisstjórnin sé stödd í miðjum hvirfilbyl út af Íslandsbankamálinu og að lítið þurfi til að upp úr slitni.	
Gullfaxi <sub>BASE</sub> (RRN)		ROUGE F1
<b>Title</b>	stjórn málafræðingur um söluna á íslandsbanka	0.00/0.00/0.00
<b>Intro</b>	eiríkur bergmann, stjórn málafræðingur og prófessor í stjórn málafræði, segir að ríkisstjórnin ætli ekki að selja hlut ríkisins í íslandsbanka. eiríkur bergmann, stjórn málafræðingur og prófessor í stjórn málafræði, segir að ríkisstjórnin ætli ekki að selja hlut sinn í íslandsbanka.	35.29/14.04/31.37
<b>Summary</b>	stjórn málafræðingur segir að salan á hlut ríkisins í íslandsbanka hafi komið sér á óvart.	29.41/11.43/29.41
Gullfaxi <sub>BASE</sub> (CNN/DailyMail)		ROUGE F1
<b>Intro</b>	stjórnarandstaðan hefur sakað ríkisstjórnina um að rifta sölu á sölunni og krafðist þess að salan á sölunni yrði lækkuð	16.00/0.00/12.00
<b>Summary</b>	stjórnarandstaðan heldur því fram að salan á hlut í adge hafi verið knúin áfram af hvirfilbyljum. ríkisstjórnin hefur verið að reyna að koma í veg fyrir sölu á hlut í ríkisstjórninni.	18.18/0.00/18.18
Gullfaxi <sub>BASE</sub> (CNN/DailyMail + RRN)		ROUGE F1
<b>Intro</b>	eiríkur bergmann stjórn málafræðingur segir tímabært að rifta sölunni á hluthafafundi í gær. hann segist hlakka til að koma í veg fyrir að salan verði afturkölluð.	21.43/6.56/21.43
<b>Summary</b>	stjórn málafræðingur segir að gagnrýni sé komið fram vegna sölu á hlut ríkisins á hlut ríkisins á hlut á hlut ríkisins í ríkinu.	34.29/10.81/34.29
Gullfaxi <sub>LARGE</sub> (RRN)		ROUGE F1
<b>Title</b>	eignatengsl á hesteyri	0.00/0.00/0.00
<b>Intro</b>	eiríkur bergmann einarsson, prófessor í stjórn málafræði við háskóla íslands og formaður sjálfstæðisflokksins, segir að ríkisstjórnin hafi ekki gripið til aðgerða til að koma í veg fyrir sölu á hlut ríkisins í íslandsbanka. eiríkur bergmann, prófessor í stjórn málafræði, segir að ríkisstjórnin hafi ekki gripið til aðgerða til að koma í veg fyrir sölu á hlut ríkisins í íslandsbanka.	31.25/16.44/28.12
<b>Summary</b>	formaður sjálfstæðisflokksins segir að salan á hlut ríkisins í íslandsbanka hafi ekki verið meiri.	22.86/5.71/22.86.
mT5 <sub>BASE</sub> (RRN)		ROUGE F1
<b>Title</b>	Þór um stjórnarsamstarf	0.00/0.00/0.00
<b>Intro</b>	Eiríkur Bergmann stjórn málafræðingur segir að flækjustigið við söluna á Íslandsbanka muni reyna enn frekar á stjórnarsamstarfið.	40.82/30.77/40.82
<b>Summary</b>	Sjálf málafræðingur segir að sú gagnrýni sem komið hefur fram vegna sölunnar á Íslandsbanka muni reyna enn frekar á stjórnarsamstarfið. Ríkisstjórnin sé nú stödd í miðjum hvirfilbyl og að lítið þurfi til að slitni upp úr.	60.38/39.29/56.60

---

**RRN Document**

---

<b>Title</b>	Ríkisstjórnin breytir samkomutakmörkunum
<b>Intro</b>	Heilbrigðisráðherra hefur fallist á tillögur sóttvarnalækis um óbreyttar sóttvarnareglur innanlands. Þá hefur verið ákveðið að bólusetningavottorð komufarþega utan Schengen-svæðisins verði tekin gild.
<b>Main</b>	Tillögur sóttvarnalækis um sóttvarnaaðgerðir innanlands voru samþykktar á ríkisstjórnarfundum í morgun. Fyrirkomulagið verður að mestu óbreytt næstu þrjár vikurnar en skerpt hefur verið á ýmsum reglum. Eins og til að mynda það að fólk haldi kyrru fyrir í hléum á tónleikum og í leikhúsum og svo framvegis og aðeins skerpt á málum varðandi hlaðborð. Ég held að fólk sé að hugsa um ferminngarlotuna sem fram undan er. Þetta eru ekki breytingar sem skipta neinu stóru þannig séð. Aðgerðir á landamærum voru ekki til umfjöllunar á fundinum. Ég geri ráð fyrir því að við munum halda ráðherranefndarfund um þau mál á næstu dögum til að fara yfir þessar tillögur sem sóttvarnalæknir er með í minnisblaðinu sem lúta að til dæmis því að prófa börn á landamærum. Hann er að fjalla líka um aukna notkun á sóttvarnahúsi og skoða möguleika á því og svo hefur hann líka verið að ræða um bólusetningavottorð á landamærum. Þá var ákveðið að þeir sem eru bólusettir og búsettir utan Schengen-svæðisins megi koma til landsins, framvísi þeir gildu vottorði um slíkt. Ferðaþjónustan hefur þrýst á um þessar breytingar. Það er mikilvægt að taka sömu vottorð gilda sama hvort þau koma innan eða utan Evrópu og í þessu felast bara tækifæri. Eruði að horfa á einhverja ákveðna markaði? Auðvitað Bretland og Bandaríkin, okkar stærstu markaði. Það liggja tækifæri í þeim. Þau eru komin lengra í bólusetningum en Evrópa og ljóst að það sé mikilvægt að fá þau hérna inn fyrir atvinnulífið hér á landi.
<b>Summary</b>	Ákveðið var á ríkisstjórnarfundum að taka gild, bólusetningavottorð komufarþega utan Schengen svæðisins. Sóttvarnareglur innanlands verða að mestu óbreyttar næstu þrjár vikurnar.

<b>Gullfaxi<sub>BASE</sub> (RRN)</b>		<b>ROUGE F1</b>
<b>Title</b>	þórólfrur um bólusetningar	0.00/0.00/0.00
<b>Intro</b>	sóttvarnalæknir hvetur fólk til að kynna sér nýjar leiðir til að koma í veg fyrir að hægt verði að koma í veg fyrir frekari breytingar á landamærum ríkjanna.	14.63/0.00/14.63
<b>Summary</b>	sóttvarnalæknir hefur falið heilbrigðisráðherra að undirbúa tillögur um hertar sóttvarnaaðgerðir til að koma í veg fyrir frekari útbreiðslu veirunnar.	5.13/0.00/5.13

<b>Gullfaxi<sub>BASE</sub> (CNN/DailyMail)</b>		<b>ROUGE F1</b>
<b>Intro</b>	the cdc er að þrýsta á að fjalla um aukna notkun á landamærum í evrópu. það eru engar breytingar á landamærum á landamærum í evrópu.	15.79/0.00/15.79
<b>Summary</b>	á fundi í morgun, heilbrigðisráðherra hefur ákveðið að gera breytingar á landamærum landsins. á undanförunum þremur vikum, cdc varar við því að breytingar á landamærum séu nauðsynlegar til að koma í veg fyrir að fólk haldi áfram í hléum.	12.24/0.00/8.16

<b>Gullfaxi<sub>BASE</sub> (CNN/DailyMail + RRN)</b>		<b>ROUGE F1</b>
<b>Intro</b>	nauðsynlegt er að fá úr því á landamærum, segir sóttvarnalæknir í svari við símtali sóttvarnalækis um aukna notkun sóttvarna á sóttvarnahúsum. ekki er gert ráð fyrir breytingum á sóttvarnalögum á næstu dögum.	16.33/3.85/16.33
<b>Summary</b>	heilbrigðisráðherra hefur ákveðið að leggja til hertar sóttvarnareglur um sóttvarnareglur um innanlandsreglur innanlands.	25.00/0.00/25.00

<b>Gullfaxi<sub>LARGE</sub> (RRN)</b>		<b>ROUGE F1</b>
<b>Title</b>	þórólfrur um sóttvarnaaðgerðir	0.00/0.00/0.00
<b>Intro</b>	sóttvarnalæknir hvetur stjórnvöld til að grípa til aðgerða til að koma í veg fyrir frekari útbreiðslu farsóttu.	5.71/0.00/5.71
<b>Summary</b>	sóttvarnalæknir ætlar að leggja til að í næstu viku verði gefin út reglugerð í sóttvarna sóttvarnareglugerð sem tekur gildi í næstu viku.	10.53/0.00/10.53.

<b>mT5<sub>BASE</sub> (RRN)</b>		<b>ROUGE F1</b>
<b>Title</b>	Bólusetningavottorð	0.00/0.00/0.00
<b>Intro</b>	Ferðaþjónustan hefur þrýst á um þessar tillögur sóttvarnalækis um sóttvarnaaðgerðir innanlands. Fyrirkomulagið verður að mestu óbreytt næstu þrjár vikurnar.	35.90/10.26/35.90
<b>Summary</b>	Heilbrigðisráðherra hefur fallist á tillögur sóttvarnalækis um óbreyttar sóttvarnareglur innanlands. Þá hefur verið ákveðið að bólusetningavottorð komufarþega utan Schengen-svæðisins verði tekin gild.	43.90/14.63/39.02

---

### C. Example Model Outputs

RRN Document		
<b>Title</b>	Björgunaraðgerðir á jökli	
<b>Intro</b>	Fyrsti hluti fjórtán manna hóps sem lenti í hrakningu á Vatnajökli í gær er kominn í fjöldahjálparstöð á Höfn í Hornafirði. Fólkið var hætt komið á jöklinum og þurfti að vera þar í sólarhring áður en það komst til byggða. Síðustu ferðamennirnir í hópnum eru væntanlegir til Hafnar eftir nokkra klukkutíma.	
<b>Main</b>	Þrjár af fjórtán manna hópi erlendra ferðamanna sem lenti í hrakningum á Hvannadalshnjúk í Vatnajökli í gær komu til Hafnar í Hornafirði nú í hádeginu. Von er á fjórum til viðbótar eftir um klukkustund. Fólkið fær aðhlyningu í fjöldahjálparstöð. Björgunarsveitarfólk á vélsleðum beið með hópnum á bálhvössu vedri og skafrenningi jöklinum í fimm klukkutíma í nótt eftir að jeppar kæmst á staðinn til að ferja fólkið niður. Björgunarmaður í aðgerðastjórn segir að fólkinu hafi verið orðið mjög kalt og ástandið á því orðið tvísýnt þegar fyrsta björgunarfólkið fann þau. Tólf ferðamenn frá Póllandi, ásamt tveimur íslenskum leiðsögumönnum, voru á leið niður af Hvannadalshnjúk síðdegis í gær þegar gps-búnaður þeirra bilaði. Skyggni var mjög slæmt. Hópurinn hélt því kyrru fyrir upp við öskjuna í Öræfajökli og kallaði eftir aðstoð björgunarsveita klukkan fjögur. Jens Olsen, varaformaður Björgunarfélags Hornafjarðar, er í aðgerðastjórn. Við sendum fjóra sleða strax af stað héðan. Svo þegar við sáum í hvað stefndi, hvernig færið var á jöklinum, þá kölluðum við út back-up. Kallað var eftir björgunarfólk á vélsleðum og bílum allt frá höfuðborgarsvæðinu austur á firði. Björgunarfólk fór upp á þremur stöðum, upp Tungnárjökul, Skálafellsjökul og frá Snæfelli. Fyrsti hópurinn kom að fólkinu um klukkan ellefu í gærkvöldi. Það voru náttúrulega bara fjórir sleðar. Það var ákveðið að þeir myndu halda kyrru fyrir og gefa fólkinu eitthvað að japla og drekka og bíða eftir fleiri sleðum. Þeir voru bara á jöklinum uppi í öskjunni í 1800 metra hæð og voru það í tveimur tjöldum. - Hvernig var vedrið þarna uppi þá? Það var svona tólf sextán metrar í vind og skafrenningur á köflum. Skyggnið var ekki neitt og rosa erfitt að gera grein fyrir ójöfnu og öðru á jöklinum. Mjög þungfært er á jöklinum og á leiðinni niður af jöklinum er sprunga farin að myndast. Um þrjúleytið bættust fleiri vélsleðar á staðinn. Bílar voru komnir nógu nálægt um fimmleytið í morgun til þess að unnt væri að ferja fólkið á sleðum í bílana. Engin börn er í hópnum. En var fólkið hætt komið á jöklinum? Samkvæmt því sem ég hef heyrt þá voru margir þarna orðnir mjög kaldir. Það var að fara að verða kítískt ástand þarna uppi. - Var fólkið orðið skelkað? Já, ég held það. Það var náttúrulega búíð að dúsa þarna í sólarhring áður en það var farið að flytja það til baka. Það lagði af stað klukkan þrjú nóttina áður í gönguna. Rauði krossinn opnaði fjöldahjálparstöð fyrir hópinn og slysavarnarfélagið eldaði kjötsúpu. Fyrsti hópurinn kom til Hafnar í hádeginu en tveir tímar þrjár tímar gætu verið í næsta hóp. Aðgerðin hefur verið mjög umfangsmikil, staðið í hátt í sólarhring. Hundrað og fjórutíu hafa tekið þátt í björgunaraðgerðinni.	
<b>Summary</b>	Fyrsti hluti fjórtán manna hóps sem lenti í hrakningu á Vatnajökli í gær er kominn í fjöldahjálparstöð á Höfn í Hornafirði. Fólkið var hætt komið á jöklinum og þurfti að vera þar í sólarhring áður en það komst til byggða. Síðustu ferðamennirnir í hópnum eru væntanlegir til Hafnar eftir nokkra klukkutíma.	
Gullfaxi <sub>BASE</sub> (RRN)		ROUGE F1
<b>Title</b>	ferðamennirnir á höfn í hornafirði	25.00/0.00/25.00
<b>Intro</b>	björgunarsveitarmenn á höfn í hornafirði voru í óða önn að bregðast við á jöklinum í gær. björgunarsveitarmenn voru í viðbragðsstöðu í jöklinum í gær.	24.56/14.29/14.04
<b>Summary</b>	tveir erlendir ferðamenn voru í fjöldahjálparstöð á höfn í hornafirði í gær, þar á meðal erlendir ferðamenn.	24.56/14.93/21.05
Gullfaxi <sub>BASE</sub> (CNN/DailyMail)		ROUGE F1
<b>Intro</b>	hópur erlendra ferðamanna lenti í vandræðum í gær í jökli í gær. fimm manns voru fluttir á jökulinn í gær eftir að hafa verið fluttir á sjúkrahús.	19.35/5.48/16.13
<b>Summary</b>	ferðamennirnir voru á jökli þegar þyrlla þeirra lenti á jökli í gær	18.52/3.28/18.52
Gullfaxi <sub>BASE</sub> (CNN/DailyMail + RRN)		ROUGE F1
<b>Intro</b>	fólk sem lenti í jöklinum á jöklinum í gær var kallað á hjálp í fjöldahjálparstöð. björgunarsveitarfólk var á meðal þeirra sem bjargað var úr hópnum.	30.00/13.51/26.67
<b>Summary</b>	fyrsti hópur manna lenti í gær á vatnajökli í gær. fólkið var hætt á jöklinum og þurfti að vera þar í nokkra klukkutíma áður en það komst til hafnar.	71.43/48.10/68.57
Gullfaxi <sub>LARGE</sub> (RRN)		ROUGE F1
<b>Title</b>	björgunarsveitarmenn á höfn í hornafirði	25.00/0.00/25.00
<b>Intro</b>	björgunarsveitir á suðurlandi voru kallaðar út í nótt til að aðstoða ferðamenn á jöklinum. björgunarsveitarmenn voru kallaðir út á sjötta tímanum í gærkvöld.	16.39/2.78/9.84
<b>Summary</b>	björgunarsveitir voru kallaðar út í morgun til að leita að erlendum ferðamönnum sem voru á leið á vatnajökul.	20.00/0.00/13.33.
mT5 <sub>BASE</sub> (RRN)		ROUGE F1
<b>Title</b>	Staðan á Höfn	33.33/0.00/33.33
<b>Intro</b>	Þrjár af fjórtán manna hópi erlendra ferðamanna sem lenti í hrakningum á Hvannadalshnjúk í gærkvöldi komu til Hafnar í Hornafirði. Von er á fjórum til viðbótar eftir um klukkustund.	32.35/12.82/32.35
<b>Summary</b>	Fyrsti hluti fjórtán manna hóps sem lenti í hrakningu á Vatnajökli í gær er kominn í fjöldahjálparstöð á Höfn í Hornafirði. Fólkið var hætt komið á jöklinum og þurfti að vera þar í sólarhring áður en það komst til byggða.	86.08/81.32/86.08

RRN Document

<b>Title</b>	Íþróttir	
<b>Intro</b>	Karlalandslið Íslands í körfubolta sækir Ítalíu heim í undankeppni heimsmeistaramótsins í kvöld. Landsliðsmaðurinn Martin Hermannsson býst við árásgjörnum leikmönnum Ítalíu í leiknum.	
<b>Main</b>	Þrjú efstu liðin í riðlinum fara áfram í milliriðil og taka þangað með sér stigin úr þessum riðli. Ítalir eiga harma að hefna eftir tapið á Íslandi á fimmtudag og þurfa nauðsynlega á sigri að halda. Landsliðsmaðurinn Martin Hermannsson býst við erfiðum leik í kvöld en segir sigur þó mögulegan. sagði Martin Hermannsson. Leikur Íslands og Ítalíu er klukkan hálfátta í kvöld og verður sýndur beint á RÚV 2, útsending hefst í HM-stofunni klukkan 19:10. Búið er að fresta leik Hollands og Rússlands í sama riðli sem átti að fara fram í kvöld. Hollendingar neita að taka á móti Rússum og sagði Hannes Jónsson, formaður körfuknattleikssambands Íslands að miðað við samtöl hans við evrópsku körfuboltaahreyfinguna þá vilji flestir að Rússum verði vikið úr undankeppninni. Þeir eru efstir í riðlinum og munu lenda í riðli með Úkraínu á næsta stigi undankeppninnar. Rússinn Roman Abramovich, eigandi enska knattspyrnufélagsins Chelsea, tilkynnti í gærkvöld að hann hafi stigið til hliðar sem stjórnandi félagsins. Abramovich útskýrir ákvörðunina ekki nánar í yfirlýsingunni sem hann gefur út tveimur dögum eftir innrás Rússa í Úkraínu. Abramovich sem keypti Chelsea árið 2003 er einn ríkasti maður Rússlands og hefur oft verið sagður náinn vinur Vladímírs Pútíns Rússlandsforseta. Manchester City endurheimti sex stiga forystu á toppi ensku úrvalsdeildarinnar í fótbolta í gærkvöld með 1-0 sigri á Everton. Phil Foden skoraði sigurmarkið undir lok leiksins. Liverpool er í 2. sæti og á leik til góða en mætir Chelsea í úrlitaleik ensku deildarbikarkeppninnar í dag. Leeds United rak í morgun knattspyrnustjóran Marcelo Bielsa. Allt hefur gengið á afturfótunum hjá Leeds að undanförunu og liðið tapaði 4-0 fyrir Tottenham í gær, hefur fengið á sig 17 mörk í síðustu fjórum leikjum og er tveimur stigum frá fallsæti	
<b>Summary</b>	Martin Hermannsson, landsliðsmaður í körfubolta, býst við grimnum leikmönnum Ítalíu þegar Ísland sækir Ítala heim í undankeppni heimsmeistaramótsins í kvöld. Ítalir þurfa nauðsynlega á sigri að halda eftir tapið á Íslandi á fimmtudag.	
<b>Gullfaxi<sub>BASE</sub> (RRN)</b>		<b>ROUGE F1</b>
<b>Title</b>	Íþróttir	100.00/0.00/100.00
<b>Intro</b>	Íslenska karlalandsliðið í fótbolta mætir Ítalíu í undankeppni heimsmeistaramótsins í fótbolta í kvöld.	37.04/31.25/37.04
<b>Summary</b>	karlalandslið Íslands í körfubolta mætir Ítalíu í undankeppni evrópumótsins í körfubolta í kvöld.	20.51/9.09/20.51
<b>Gullfaxi<sub>BASE</sub> (CNN/DailyMail)</b>		<b>ROUGE F1</b>
<b>Intro</b>	hollendingar töpuðu fyrir Íslandi í undankeppni em í knattspyrnu á fimmtudag. hollendingar töpuðu fyrir Rússum í opunarleik d-riðils í undankeppni em. hollendingar eru í riðli í umspili um sæti í undankeppni em í knattspyrnu.	11.11/4.35/11.11
<b>Summary</b>	martin, landsliðsþjálfari Íslands, býst við erfiðum leik í undankeppni em í kvöld. hollendingar töpuðu 1-0 fyrir Þjóðverjum í opunarleik d-riðils í undankeppni em í d-riðli.	20.83/11.32/20.83
<b>Gullfaxi<sub>BASE</sub> (CNN/DailyMail + RRN)</b>		<b>ROUGE F1</b>
<b>Intro</b>	Ítalska kvennalandsliðið í körfubolta mætir Ítalíu í lokaleik sínum í undankeppni heimsmeistaramótsins í körfubolta í kvöld. Rússar eru með sex stiga forskot á toppi deildarinnar.	31.58/27.27/31.58
<b>Summary</b>	karlalandslið Íslands í körfubolta sækir í undankeppni heimsmeistaramótsins í körfubolta í kvöld. karlalandsliðið í körfubolta býst við erfiðum leik.	33.33/20.83/33.33
<b>Gullfaxi<sub>LARGE</sub> (RRN)</b>		<b>ROUGE F1</b>
<b>Title</b>	Íþróttir	100.00/0.00/100.00
<b>Intro</b>	Íslenska karlalandsliðið í körfubolta mætir Ítalíu í lokaleik sínum í undankeppni hm í körfubolta í kvöld. landsliðsþjálfari Íslands segist ekki hafa áhyggjur af því.	32.43/18.60/32.43
<b>Summary</b>	Íslenska karlalandsliðið í körfubolta mætir Slóvakíu í undankeppni em í dag.	10.26/4.65/10.26.
<b>mT5<sub>BASE</sub> (RRN)</b>		<b>ROUGE F1</b>
<b>Title</b>	Íþróttir	100.00/0.00/100.00
<b>Intro</b>	Rússneski knattspyrnumaðurinn Martin Hermannsson býst við erfiðum leik í kvöld en segir sigur þó mögulegan.	36.36/22.86/36.36
<b>Summary</b>	Karlalandslið Íslands í körfubolta sækir Ítalíu heim í undankeppni heimsmeistaramótsins í kvöld. Landsliðsmaðurinn Martin Hermannsson býst við árásgjörnum leikmönnum Ítalíu í leiknum.	46.81/26.42/46.81



## C. Example Model Outputs

RRN Document		
<b>Title</b>	Alþjóðleg jarðhitaráðstefna	
<b>Intro</b>	Vonir standa til þess að nýr sjálfbærnistaðall, þróaður af Íslendingum, auðveldi fjármögnun nýrra jarðhitaverkefna víða um heim. Heimsþing Alþjóða jarðhitasambandsins hefst í Hörpu í dag og von er á 1100 gestum til að ræða það nýjasta í nýtingu jarðvarma.	
<b>Main</b>	Heimsþing Alþjóða jarðhitasambandsins er haldið á fimm ára fresti. Stærðarlega séð er ráðstefnan á pari við Hringborð Norðurslóða. Auk 1100 gesta í raunheimum er gert ráð fyrir annarri eins þátttöku í netheimum. Þá ætla 250 sérfræðingar frá 100 löndum að deila þekkingu sinni á jarðvarma. Bjarni Pálsson, formaður undirbúningsnefndar Íslands á ráðstefnunni, segir ýmislegt hafa gerst í jarðhitamálum frá síðustu þingi. Það má segja að vaxtarsprotarnir séu víða, það er bæði verið að færa jarðhitanytingu til fleiri landa sem ekki hafa verið að nýta jarðhita áður, það er mikil uppbygging að eiga sér stað til dæmis í Austur-Afríku. Þau lönd sem hafa meiri sögu og reynslu af því að nýta jarðvarma eru bæði að færa sig neðar, dýpra í jörðina í meiri orku og einnig er verið að leita meira í lægri hita, bæta nýtingu á svæðum þar sem er kannski hiti sem áður hefði ekki talist nægilega fýsilegur til að nýta. Það er að eiga sér stað víða um Evrópu, til dæmis í gömlum kolanámum og olíuborholum. Eitt af því sem helst hefur staðið uppbyggingu jarðhita fyrir þrífum, að sögn Bjarna, er tregða fjárfesta við að taka kostnaðarsama áhættu. Hann telur að nýr sjálfbærnistaðall, sem hefur verið þróaður hér út frá sambærilegum, alþjóðlegum staðli fyrir vatnsaflsvirkjanir, geti verið liður í því að treysta fjármögnun til framtíðar og ná breiðari sátt meðal umhverfissamtaka, fjárfesta og fleiri um nýtingu jarðhita. Þá geti náðst mælikvarði á hvort verkefni séu hæf til að fá svokallaða græna fjármögnun. Það er auðvitað töl sem mörg stjórnvöld eru að reyna að vinna að núna, hvernig eigi að meta góð verkefni sem styðja við aðgerðir gegn loftslagsbreytingum og þar með hjálpa þessum verkefnum í þeirra samkeppni við aðra orkugjafa sem eru ekki jafn umhverfisvænir	
<b>Summary</b>	Vonir standa til þess að nýr sjálfbærnistaðall, þróaður af Íslendingum, auðveldi fjármögnun nýrra jarðhitaverkefna víða um heim. Yfir tvöþúsund manns taka þátt í alþjóðlegri jarðhitaráðstefnu sem hefst í Hörpu í dag.	
Gullfaxi <sub>BASE</sub> (RRN)		ROUGE F1
<b>Title</b>	jarðhitapíng	0.00/0.00/0.00
<b>Intro</b>	ráðstefna sameinuðu þjóðanna um nýtingu jarðhitans í heiminum hefst í dag. í henni er meðal annars lögð áhersla á nýtingu endurnýjanlegra orkulinda og nýtingu endurnýjanlegra orkulinda.	29.09/6.56/21.82
<b>Summary</b>	vonir standa til að hægt verði að hrinda í framkvæmd alþjóðlegri orkuráðstefnu sameinuðu þjóðanna sem haldin verður í hörpu í dag.	40.82/23.08/40.82
Gullfaxi <sub>BASE</sub> (CNN/DailyMail)		ROUGE F1
<b>Intro</b>	ráðstefnan er haldin í hollandi á fimm ára fresti. ráðstefnan er gert ráð fyrir að halda áfram í þessari viku.	15.38/0.00/11.54
<b>Summary</b>	á ráðstefnu í dag er gert ráð fyrir að ræða stærsta jarðhitaverkefni heims. ráðstefnan er í gangi á fimm ára fresti til að ræða jarðhitanytingu.	16.00/3.70/16.00
Gullfaxi <sub>BASE</sub> (CNN/DailyMail + RRN)		ROUGE F1
<b>Intro</b>	á heimsmeistaramótinu í nýtingunýtingu jarðhita verður lýst á alþingi í dag. ráðstefnan er hafin til að nýta betur þá sem nýta endurnýjanlega orku.	21.82/6.67/18.18
<b>Summary</b>	vonir standa til þess að nýr staðall nýrra jarðhitaverkefna verði opnaður í hörpu í dag.	53.33/43.48/53.33
Gullfaxi <sub>LARGE</sub> (RRN)		ROUGE F1
<b>Title</b>	jarðhitaráðstefna um jarðhitanytingu	40.00/0.00/40.00
<b>Intro</b>	framkvæmdastjóri sameinuðu þjóðanna segir að stefnt sé að því að hrinda í framkvæmd hugmyndum um nýtingu endurnýjanlegra orkugjafa.	15.69/0.00/15.69
<b>Summary</b>	stefnt er að því að hrinda í framkvæmd alþjóðlegri hugmyndasamkeppni um jarðhitasvæði á íslandi í næstu viku.	21.74/0.00/21.74.
mT5 <sub>BASE</sub> (RRN)		ROUGE F1
<b>Title</b>	Jarðhitaráðstefna	66.67/0.00/66.67
<b>Intro</b>	Formaður undirbúningsnefndar Íslands á ráðstefnu Alþjóða jarðhitasambandsins segir mikil uppbygging að eiga sér stað víða um Evrópu. Formaður undirbúningsnefndar Íslands telur að nýtingu jarðhita geti verið liður í því að treysta fjárfesta við að taka kostnaðarsama áhættu.	24.24/5.56/18.18
<b>Summary</b>	Viðræður standa til þess að nýr sjálfbærnistaðall, þróaður af Íslendingum, auðveldi fjármögnun nýrra jarðhitaverkefna víða um heim. Vonir standa til þess að nýr sjálfbærnistaðall, þróaður af Íslendingum, auðveldi fjármögnun nýrra jarðhitaverkefna víða um heim.	70.83/65.31/70.83





---

**RRN Document**

---

<b>Title</b>	Vill sérlög um Seyðisfjörð
<b>Intro</b>	Forseti sveitarstjórnar Múlaþings vill að Alþingi setji sérlög um Seyðisfjörð til að bregðast við ósanngirni í reglum Ofanflóðasjóðs. Atvinnuhúsnæði fæst ekki bætt þó atvinnulíf greiði í sjóðinn og eigendur íbúðarhúsa fá sumir minna en fasteignamat gerir ráð fyrir, fyrir húsin sín.
<b>Main</b>	Forseti Sveitastjórnar Múlaþings telur að reglur um Ofanflóðasjóð séu ósanngjarnar gagnvart eignum atvinnuhúsnæðis á Seyðisfirði. Líka gagnvart þeim sem mega ekki búa í húsum sínum. Mögulega þyrfti að setja sérlög um Seyðisfjörð. Lög og reglur varðandi Ofanflóðasjóð eru á þann hátt að íbúðarhúsnæði fæst bætt að hluta. En atvinnuhúsnæði fæst það ekki. Sem er sérkennilegt sérstaklega í ljósi þess að atvinnulíf greiðir í þennan umrædda Ofanflóðasjóð líkt og aðrir. Þannig að það eitt og sér er dálítið sérstakt. (Og hvaða lausnir vilja menn sjá á þessu til að hægt verði að flytja frystihúsið á Seyðisfirði?) Leiðin sem ég hef rætt innan samstarfshópsins hún er sú að það verði þá gripð mögulega til einhverra sértækra aðgerða á Seyðisfirði til að bregðast við vandanum sem við stöndum frammi fyrir. Það gæti mögulega þýtt lagasetning. Segir Gauti Jóhannesson, forseti sveitastjórnar Múlaþings, en hann á sæti í samráðshópi stjórnvalda vegna skriðufallanna á Seyðisfirði í desember. Eins og fram hefur komið er frystihús Síldarvinnslunnar, stærsti vinnustaðurinn í bænum á hættusvæði í nýju bráðabirgðahættumati. Þar er erfitt að rýma og spá fyrir um skriðuhættu vegna bráðnandi sífrera í Strandartindi. Altjón varð á 13 húsum á Seyðisfirði og fá eigendur þeirra greitt samkvæmt brunabótamati. Bannað er að búa í fjórum til viðbótar samkvæmt nýju hættumati og ákveðin mismunur snýr að þeim sem eiga bannhús. Þeir fá greitt samkvæmt mati sem byggist á nýlegri sölu á húsnæði á Seyðisfirði og sú upphæð nær í sumum tilvikum ekki fasteignamati sem þó er mun lægra en brunabótamatið. Mér fundist það sanngirnisaþriðið að miða allavega við það að fólk fái á lágmarki fasteignamat fyrir umræddar húseignir.
<b>Summary</b>	Forseti sveitarstjórnar Múlaþings vill að Alþingi setji sérlög um Seyðisfjörð til að bregðast við ósanngirni í reglum Ofanflóðasjóðs. Atvinnuhúsnæði fæst ekki bætt þó atvinnulíf greiði í sjóðinn og eigendur íbúðarhúsa fá sumir minna en fasteignamat fyrir húsin sín.

<b>Gullfaxi<sub>BASE</sub> (RRN)</b>		<b>ROUGE F1</b>
<b>Title</b>	á alþingi	0.00/0.00/0.00
<b>Intro</b>	forseti bæjarstjórnar á seyðisfirði ætlar að grípa til aðgerða til að draga úr hættu á skriðuföllum.	11.54/3.64/11.54
<b>Summary</b>	forseti alþýðusambands austurlands vill að reglur um atvinnuhúsnæði brjóti í bága við lög um atvinnuhúsnæði.	31.37/3.85/27.45

<b>Gullfaxi<sub>BASE</sub> (CNN/DailyMail)</b>		<b>ROUGE F1</b>
<b>Intro</b>	í stað þess að rýma þyrfti að setja lög á húseigendur í húsum sínum. í stað þess að rýma þyrfti að setja lög í stað til að bregðast við vandamálum.	18.18/10.00/18.18
<b>Summary</b>	forseti vill ýta á að bregðast við ósanngirni í fjármálum. í stað þess að bregðast við ósanngirni, það er erfitt að spá fyrir um hvað á að gera.	31.03/12.90/27.59

<b>Gullfaxi<sub>BASE</sub> (CNN/DailyMail + RRN)</b>		<b>ROUGE F1</b>
<b>Intro</b>	stefnt er að því að leggja á atvinnuhúsnæði á seyðisfirði til að bregðast við vanda líðandi stundar vegna skriðuhættu og mögulegra náttúruhamfara. forseti bæjarstjórnar segist grípa til aðgerða til að bregðast við breyttum lögum.	21.54/8.57/21.54
<b>Summary</b>	forseti sveitarstjórnarþings vill að alþingi setji sér lög um atvinnuhúsnæði til að bregðast við ósanngirni í reglugerð um atvinnuhúsnæði. fæst ekki fæst bætt í sjóðinn og eigendur íbúðarhúsa fá minna.	66.67/41.79/66.67

<b>Gullfaxi<sub>LARGE</sub> (RRN)</b>		<b>ROUGE F1</b>
<b>Title</b>	ofanflóðasjóður á seyðisfirði	0.00/0.00/0.00
<b>Intro</b>	ofanflóðasjóður hefur áhyggjur af því að ofanflóðasjóður þurfi að grípa til aðgerða til að koma í veg fyrir tjón á húsum á seyðisfirði.	14.04/3.23/14.04
<b>Summary</b>	forseti alþingis ætlar að grípa til ráðstafana til að koma í veg fyrir uppkaua á fasteignum á seyðisfirði.	22.64/3.57/22.64.

<b>mT5<sub>BASE</sub> (RRN)</b>		<b>ROUGE F1</b>
<b>Title</b>	Ofanflóðasjóður um Seyðisfirði	28.57/0.00/28.57
<b>Intro</b>	Forseti Sveitastjórnar Múlaþings telur að reglur um Ofanflóðasjóð séu ósanngjarnar gagnvart þeim sem mega ekki búa í húsum sínum.	20.69/0.00/20.69
<b>Summary</b>	Forseti sveitarstjórnar Múlaþings vill að Alþingi setji sérlög um Seyðisfjörð til að bregðast við ósanngirni í reglum Ofanflóðasjóðs. Atvinnuhúsnæði fæst ekki bætt þó atvinnulíf greiði í sjóðinn.	80.65/81.25/80.65

---



RRN Document

<b>Title</b>	Undirbúningskjörbréfanefnd - þriðji fundur
<b>Intro</b>	Allir frambjóðendurnir fimm sem dattu út við endurtalningu í Norðvesturkjördæmi hafa nú kært kosningarnar. Alls hafa sex kærur borist undirbúningskjörbréfanefnd og sú sjöunda er á leiðinni. Birgir Ármannsson, formaður nefndarinnar, segir óljóst hvenær vinnu nefndarinnar lýkur og ógjörningur að segja til um hvort kjósa þurfi aftur í Norðvesturkjördæmi.
<b>Main</b>	Nefndin fundaði í þriðja sinn í morgun. Engir gestir komu fyrir nefndina en nefndarmenn höfðu í nógu öðru að snúast við að fara yfir ágreiningsatkvæði og ýmis gögn. Svo munum við líka fara yfir verklag okkar á næstunni varðandi þessar kærur sem eru komnar fram, og taka stöðuna á gagnaöflun og öðru slíku, við erum búin að fá töluvert af gögnum nú þegar en við eigum von á meiru. Allir frambjóðendurnir fimm sem misstu jöfnunarsæti sitt eftir endurtalninguna í Norðvesturkjördæmi hyggjast kæra kosningarnar. Rósa Björk Brynjólfssdóttir, frambjóðandi Samfylkingarinnar, Guðmundur Gunnarsson úr Viðreisn, Lenya Rún Taha Karim, frambjóðandi Pírata og Hólmfríður Árnadóttir, Vinstri grænum hafa þegar kært til Alþingis. Karl Gauti Hjaltason, Miðflokksmaður hefur kært kosningarnar til lögreglu og hyggst einnig leggja fram kæru til Alþingis. Til viðbótar hafa svo borist tvær kærur frá borgurum, annar þeirra er Magnús Davíð Norðdahl, oddviti Pírata í Norðvesturkjördæmi. Birgir segir óljóst hvenær vinnu nefndarinnar ljúki. Ég myndi segja að næsta vika skæri svolítið úr um hvað við getum gert þetta fljótt. Býstu við því að það þurfi að kjósa aftur í Norðvestur? Það er engin leið að segja.
<b>Summary</b>	Allir frambjóðendurnir fimm sem endurtalning atkvæða í Norðvesturkjördæmi hafði áhrif á hafa nú kært kosningarnar. Formaður undirbúningskjörbréfanefndar segir ómögulegt að segja til um hvort telja þurfi aftur í Norðvesturkjördæmi.

	Gullfaxi <sub>BASE</sub> (RRN)	ROUGE F1
<b>Title</b>	prófkjör í norðvesturkjördæmi	0.00/0.00/0.00
<b>Intro</b>	formaður yfirkjörstjórnar í norðvesturkjördæmi segir að ekki hafi verið hægt að afgreiða kærur í norðvesturkjördæmi.	21.43/3.39/21.43
<b>Summary</b>	tveir frambjóðendur í norðvesturkjördæmi hafa verið kærðir til lögreglu í tengslum við talningu atkvæða í norðvesturkjördæmi í síðustu viku.	27.27/8.70/22.73

	Gullfaxi <sub>BASE</sub> (CNN/DailyMail)	ROUGE F1
<b>Intro</b>	nefndin í norðvesturhluta hittast í þriðja sinn í dag til að takast á við málið. fimm frambjóðendurnir misstu rétt sinn til að snúa aftur til alþingis.	25.00/0.00/15.62
<b>Summary</b>	sex kærur hafa verið lögð fram í norðvesturkjördæmi í norðvesturhluta norðvestursins í norðvesturhluta landsins. engir gestir komu fyrir nefndina í morgun.	13.33/4.26/8.89

	Gullfaxi <sub>BASE</sub> (CNN/DailyMail + RRN)	ROUGE F1
<b>Intro</b>	iðnaðar- og efnahagsráðherra ætlar í morgun að leggja fram kæru til alþingis til að leggja fram kæru til alþingis vegna endurtalningar í norðvesturkjördæmi. fimm frambjóðendur í norðvesturkjördæmi hyggjast leggja fram kæru til alþingis í dag.	19.05/2.86/19.05
<b>Summary</b>	allir fimm frambjóðendur í norðvesturkjördæmi hafa verið ákærðir fyrir kosningarnar í norðvesturkjördæmi. þeir hafa verið ákærðir í tengslum við endurtalningu og endurtalningu.	31.82/4.26/31.82

	Gullfaxi <sub>LARGE</sub> (RRN)	ROUGE F1
<b>Title</b>	prófkjör í norðvesturkjördæmi	0.00/0.00/0.00
<b>Intro</b>	formaður stjórnskipunar- og eftirlitsnefndar alþingis segir að ekki hafi verið tekin ákvörðun um að taka þátt í talningunni í norðvesturkjördæmi í dag. formaður nefndarinnar segir að ekki hafi verið tekin ákvörðun um að taka þátt í talningunni.	25.00/2.86/21.87
<b>Summary</b>	talningu atkvæða í norðvesturkjördæmi lauk í morgun eftir að yfirkjörstjórnir í norðvesturkjördæmi greindu frá því í morgun að talningu atkvæða í norðvesturkjördæmi hefði ekki verið lokið.	22.22/8.16/22.22

	mT5 <sub>BASE</sub> (RRN)	ROUGE F1
<b>Title</b>	Kosningar í NV-kjördæmi	0.00/0.00/0.00
<b>Intro</b>	Formaður nefndarinnar segir óljóst hvenær vinnu nefndarinnar ljúki. Allir frambjóðendurnir fimm sem misstu jöfnunarsæti sitt eftir endurtalningu í Norðvesturkjördæmi hyggjast kæra kosningarnar.	43.08/26.87/43.08
<b>Summary</b>	Formaður kjörbréfanefndar í Norðvesturkjördæmi segir óljóst hvenær vinnu nefndarinnar lýkur og ógjörningur að segja til um hvort kjósa þurfi aftur í Norðvesturkjördæmi. Alls hafa sex kærur borist undirbúningskjörbréfanefnd og sú sjöunda er á leiðinni.	44.07/23.33/40.68



RRN Document		
<b>Title</b>	Dánartíðni í Evrópu 2020	
<b>Intro</b>	Rúmlega 40 prósentum fleiri létust í Evrópusambands- og EFTA-ríkjum í nóvember í fyrra en fjögur árin á undan. Þá náði önnur bylgja COVID-19 hámarki. Allt árið í fyrra voru andlát á Íslandi umfram meðaltal 1,6 prósent samkvæmt tölum Eurostat.	
<b>Main</b>	Hagstofa Evrópusambandsins, Eurostat, hefur birt bráðabirgðatölur um andlát í ríkjum sambandsins og EFTA-ríkjunum Íslandi, Noregi, Liechtensten og Sviss árið 2020. Þar má glöggst sjá áhrif COVID-19 faraldursins. Í yfirliti Eurostat kemur fram að í apríl í fyrra hafi dánartíðni í Evrópulöndum verið 25 prósentum hærrí en árin 2016-19 og í hámarki annarrar bylgju í nóvember létust rúmlega 40 prósentum fleiri í ESB- og EFTA-ríkjum en fjögur árin á undan.	
<b>Summary</b>	Rúmlega 40 prósentum fleiri létust í ESB- og EFTA-ríkjum í nóvember í fyrra en fjögur árin á undan. Einu komma sex prósentum fleiri létust á Íslandi allt árið í fyrra samkvæmt tölum Eurostat, hagstofu Evrópusambandsins.	
Gullfaxi <sub>BASE</sub> (RRN)		ROUGE F1
<b>Title</b>	þjóðaratkvæðagreiðsla í esb	28.57/0.00/28.57
<b>Intro</b>	þetta kemur fram í nýrri skýrslu eurostat, sem birt er á vef eurostat.	12.77/0.00/12.77
<b>Summary</b>	rúmlega sjötíu af hverjum tíu ríkjum evrópusambandsins létust á em í fyrra, samkvæmt nýjum tölum frá eurostat.	31.11/4.17/26.67
Gullfaxi <sub>BASE</sub> (CNN/DailyMail)		ROUGE F1
<b>Intro</b>	í apríl á þessu ári voru rúmlega 40 prósent í evrópusambandinu. í apríl á síðasta ári voru tæplega 40 prósent í evrópusambandinu. í apríl á þessu ári var talan rúmlega 40 prósentum hærrí en árið í fyrra.	37.74/13.11/30.19
<b>Summary</b>	í nóvember á síðasta ári voru 40 prósent fleiri en í nóvember á síðasta ári en í nóvember á þessu ári	31.58/4.55/31.58
Gullfaxi <sub>BASE</sub> (CNN/DailyMail + RRN)		ROUGE F1
<b>Intro</b>	í apríl í fyrra létust um 40 prósent fleiri í evrópu árið 2016 en á sama tíma í fyrra.	36.73/3.70/20.41
<b>Summary</b>	rúmlega 40 prósent fleiri létust í nóvember síðastliðnum en á síðasta ári.	40.00/18.60/40.00
Gullfaxi <sub>LARGE</sub> (RRN)		ROUGE F1
<b>Title</b>	landsframleiðsla á mann í esb	22.22/0.00/22.22
<b>Intro</b>	landsframleiðsla á mann dróst saman um 0,3% á síðasta ári og jókst um 0,3% á milli ára. þetta kemur fram í nýrri skýrslu eurostat.	14.81/0.00/14.81
<b>Summary</b>	á fyrsta fjórðungi þessa árs var ísland í efsta sæti í framkvæmdastjórn evrópusambandsins en á sama tíma í fyrra var ísland í efsta sæti.	23.26/4.00/18.60.
mT5 <sub>BASE</sub> (RRN)		ROUGE F1
<b>Title</b>	Staðan í Evrópusambands- og EFTA-ríkjum	22.22/0.00/22.22
<b>Intro</b>	Forsætisráðherra Evrópusambandsins segir bráðabirgðatölur um andlát í ríkjum sambandsins og EFTA-ríkjum Evrópusambandsins hafa sýnt áhrif faraldursins.	16.33/3.85/16.33
<b>Summary</b>	Rúmlega 40 prósentum fleiri létust í Evrópusambands- og EFTA-ríkjum í nóvember í fyrra en fjögur árin á undan. Þá náði önnur bylgja COVID-19 hámarki.	61.22/55.56/61.22





# D. Human Evaluation Examples

Example outputs and their human evaluation for the Main  $\rightarrow$  Intro task. The annotations listed are relevance, correctness and language (R/C/L) as defined in Section 3.4.2.

RRN Document		
<b>Title</b>	Íþróttir	
<b>Intro</b>	Valur tryggði sér í gærkvöld Íslandsmeistaratitilinn í körfubolta karla eftir sigur á Tindastóli í oddaleik úrslitaeinvígisins. Það er skammt stórra högga á milli hjá Valsmönnum því næsta úrslitaeinvígi hefst í kvöld.	
<b>Main</b>	Tindastóll byrjaði leikinn betur og var með átta stiga forystu eftir fyrsta leikhluta. Valsmenn léku betur í öðrum leikhluta og þeir komust í fyrsta inn yfir í leiknum í stöðunni 32-30. Í hálfleik var staðan hnífjöfn, 36-36. Taugatrekkingurinn var mikill í seinni hálfleik og það sást á báðum liðum. Fyrir lokafjórðunginn munaði einu stigi, 50-49, en Valur kláraði svo einvígið með frábærum endasprið og vann með þrettán stiga mun, 73-60. Þetta er fyrsti Íslandsmeistaratitill Valsmanna í 39 ár eða frá árinu 1983. sagði Helgi Rafn Viggósson, leikmaður Tindastóls, en áður var rætt við nýkrýndan Íslandsmeistara Kristófer Acox. Valsarar þurfa ekki að bíða lengi eftir næsta úrslitaeinvígi því það hefst strax í kvöld. Valur og ÍBV hefja þá einvígið um Íslandsmeistaratitil karla í handbolta og fer fyrsti leikurinn fram á Hlíðarenda klukkan hálf átta. Líkt og í körfunni þarf að vinna þrjá leiki til að tryggja Íslandsmeistaratitilinn. Annað kvöld er svo komið að einvígi Vals og Fram um Íslandsmeistaratitil kvenna í handbolta og því ljóst að stuðningsmenn Vals hafa í nógu að snúast þessa dagana.	
<b>Summary</b>	Valur vann í gærkvöld sinn fyrsta Íslandsmeistaratitil í körfubolta karla í 39 ár eftir sigur á Tindastóli í fimmta leik úrslitaeinvígisins.	
Gullfaxi <sub>BASE</sub> (RRN)		R/C/L
<b>Intro</b>	tindastóll og valur mætast í úrvalsdeild karla í körfubolta í kvöld.	1/0/1
Gullfaxi <sub>BASE</sub> (CNN/DailyMail)		R/C/L
<b>Intro</b>	a fullur-bakvörður lék í seinni hálfleik í leik liðanna í 32 - 0 sigri á heimavelli í fyrri leik liðanna í 32 liða úrslitum evrópumótsins í kvöld	0/0/0
Gullfaxi <sub>BASE</sub> (CNN/DailyMail + RRN)		R/C/L
<b>Intro</b>	valsmenn eru komnir í 2-0 í úrslitaeinvíginu um íslandsmeistaratitil karla í körfubolta eftir sigur á haukum í öðrum leikhluta í gærkvöld.	1/0/1
mT5 <sub>BASE</sub> (RRN)		R/C/L
<b>Intro</b>	átta liða úrslitaeinvígið um íslandsmeistaratitilinn í handbolta hefst í kvöld. þetta er fyrsti íslandsmeistaratitil valsmanna í 39 ár.	1/0/0

## D. Human Evaluation Examples

RRN Document		
<b>Title</b>	Deildarmyrkvi	
<b>Intro</b>	Skammdegishimininn yfir landinu bauð upp á mikið sjónarspil í morgun þegar jörðin varpaði skugga sínum á tunglið svo úr varð svokallaður deildarmyrkvi. Myrkvinn hófst þegar klukkan var nítján mínútur gengin í átta og náði hámarki um þrjár mínútur yfir níu. Þegar mest var þakti skuggi jarðar um 97 prósent af tunglinu. Um stund var tunglið baðað rauðum bjarma sem vel mátti sjá með berum augum. Sævar Helgi Bragason ritstjóri Stjórnfræðivefsins segir að rauða litinn megi rekja til sólarljóss sem berst í gegnum andrúmsloft jarðar.	
<b>Main</b>	Svona myrkvar eiga sér stað tvisvar til þrisvar á ári. Við þurfum bara að vera á réttum stað á réttum tíma til þess að sjá hann. Það er algengara að við fáum svona deildarmyrkva frekar en almyrkva. Þannig að þetta gerist annað slag. Þetta er ekkert mjög sjaldgæft en á tilteknum stöðum á jörðinni ekkert rosalega oft hins vegar. Er þetta svipað eins og sólmyrkvi jafnvel? Sólmyrkvar eru tilkomumeiri sérstaklega almyrkvar. Sólmyrkvar eru aðeins sjaldséðari, það er að segja svona almyrkvar, heldur en almyrkvar á tungli. Það er reyndar einn fram undan 4. desember á Suðurskautinu. Ég ætlaði að vera þar en það var eitt og annað sem kom í veg fyrir það. Þannig að því miður fæ ég ekki að njóta þess að þessu sinni.	
<b>Summary</b>		
<b>Gullfaxi<sub>BASE</sub> (RRN)</b>		
<b>Intro</b>	myrkvinn á sólu í dag er sá stærsti í heiminum frá því í fyrra. myrkvi á sólu er sá heitasti í heiminum.	R/C/L 1/0/0
<b>Gullfaxi<sub>BASE</sub> (CNN/DailyMail)</b>		
<b>Intro</b>	þetta er í annað sinn sem myrkvi hefur verið á jörðinni á þessu ári. þetta er í annað sinn sem það gerist í desember á þessu ári.	R/C/L 1/0/0
<b>Gullfaxi<sub>BASE</sub> (CNN/DailyMail + RRN)</b>		
<b>Intro</b>	enn er ekki vitað hvað gerist á norðurskautinu í lok myndbandsins en myrkvi á sér stað í kvöld. myrkvi á sér stað í fyrsta skipti í rúman áratug.	R/C/L 1/0/0
<b>mT5<sub>BASE</sub> (RRN)</b>		
<b>Intro</b>	forstjóri almannavarnastofnunarinnar segir að sólmyrkvar á suðurskautinu geti orðið sjaldgæft. sólmyrkvar eiga sér stað tvisvar til þrisvar á ári.	R/C/L 1/0/1

---

**RRN Document**

---

<b>Title</b>	Pórolfur um sóttvarnaaðgerðir
<b>Intro</b>	Engin áform eru um afléttingar sóttvarnaaðgerða innanlands á næstunni segir sóttvarnalæknir. Hann segir að ekki megi slaka á í sóttvörnum á landamærunum, til þess að svo megi verða þurfi liðsinni Íslenskrar erfðagreiningar við sýnatöku þar.
<b>Main</b>	Tveir greindust með kórónuveiruna í gær, báðir í sóttkví. Síðast greindist smit utan sóttkvíar 12. maí. Þórolfur Guðnason sóttvarnalæknir segir að sá fjöldi smita sem greinst hafi síðustu daga vekur vonir um að komist hafi verið fyrir þau hópsmit sem hafa verið að greinast undanfarið. Hann segir að ekki sé tilefni til að aflétta aðgerðum innanlands að svo stöddu, en fjöldi ferðamanna hafi aukist meira en búist var við. Auðvitað eru margir sem gleðjast yfir því. En það er mikil áskorun fyrir okkur að viðhafa fullnægjandi sóttvarnir á landamærunum meðan við erum að ná meiri útbreiðslu í bólusetningarnar. Telurðu ástæðu til að breyta sóttvörnum á landamærunum? Við erum búin að ná ansi góðu fyrirkomulagi hvað varðar sóttvarnir á landamærunum og ég held að það sé nokkurn veginn það öruggasta sem við getum haft núna. Ég held að það sé mikilvægt að halda því áfram þangað til við náum góðu marki í því. Þórolfur segir ljóst að til þess að viðhalda sóttvörnum á landamærunum þurfi fleiri að koma að skimunum þar. Íslensk erfðagreining skimaði við landamærin í fyrRASumar en hann segir engar formlegar viðræður í gangi um framhald á því verkefni. Ef við ætlum að halda áfram með þessar PCR-skimanir sem eru bestu greiningarnar sem við höfum þá þurfum við að fá utankomandi aðstoð. ER einhver annar sem kemur til greina en IE? Það er enginn annar aðili hér innanlands sem hefur yfir þeirri getu að ráða eins og Íslensk erfðagreining.

**Summary**

<b>Gullfaxi<sub>BASE</sub> (RRN)</b>		<b>R/C/L</b>
<b>Intro</b>	Þórolfur Guðnason, sóttvarnalæknir, segir að ekki sé hægt að útiloka að smit berist á landamærunum. Þórolfur Guðnason sóttvarnalæknir segir að ekki sé hægt að útiloka að smit berist á landamærunum.	1/1/0
<b>Gullfaxi<sub>BASE</sub> (CNN/DailyMail)</b>		<b>R/C/L</b>
<b>Intro</b>	Líkurnar á að smitast eru meira en tvöfalt meiri en búist var við. Líkurnar á að smitast af veirunni hafa aukist meira en búist var við.	1/0/0
<b>Gullfaxi<sub>BASE</sub> (CNN/DailyMail + RRN)</b>		<b>R/C/L</b>
<b>Intro</b>	Þórolfur Guðnason sóttvarnalæknir segir að ekki sé búið að sækja um kórónuveiruna hér á landi að sögn sóttvarnalæknis. Engar formlegar viðræður hafi verið í gangi um sóttvarnir hér á landi.	1/0/0
<b>mT5<sub>BASE</sub> (RRN)</b>		<b>R/C/L</b>
<b>Intro</b>	Þórolfur Guðnason sóttvarnalæknir segir að fjöldi ferðamanna hafi aukist meira en búist var við.	1/1/1

## D. Human Evaluation Examples

RRN Document		
<b>Title</b>	Örmagna framlínustarfsfólk	
<b>Intro</b>	Framlínustarfsmenn, sem staðið hafa lengi í broddi fylkingar í baráttunni við COVID-19 hér á landi, eiga á hættu að örmagnast andlega og kulna í starfi.	
<b>Main</b>	<p>Þetta segir Sandra B. Franks, formaður Sjúkraliðafélags Íslands, í aðsendri grein í Morgunblaðinu. Sjálf bauð hún sig fram í bakvarðasveit heilbrigðiskerfisins þegar faraldurinn stóð sem hæst. Álagið hafi komið fram með ýmsu móti, til að mynda í minni nánd milli sjúklinga og starfsmanna sem voru grímuklæddir í óþægilegum hlífðarbúnaði. Ég hef áhyggjur af að það sé ekki tekið nægilega vel utan um þessa starfsmenn sem eru búnir að vera lengi í framlínu. Það er ástæða fyrir því af því að ég upplifði það á eigin skinni hvernig þessu fólki leið þegar ég var að vinna þarna, og það var bara í stuttan tíma. En þetta fólk sem þarna er að vinna er að vinna með þetta veika fólk árum saman og þá uppsafnast ákveðin þreyta og óþol gagnvart því umhverfi sem það er búíð að vera í og það er loksins að sjá fyrir endann á þessu álagi sem er búíð að vera en það sem blasir við er áframhaldandi starf í þessu umhverfi. Þá hafi fólk sífellt óttast að bera smit inn á deildir, inn á eigin heimili og á milli samstarfsfólks. Fólk sem var í þessari sjálfskipuðu sóttkví var ekki að hitta fjölskylduna sína og börnin sín af því að það þurfti jafnvel að gista á hóteli á milli vaktanna. Það varð til þess að álagið inni á heimilunum sem jafnan dreifðist á alla heimilismedlimi, þá var það í höndum makans.</p>	
<b>Summary</b>		
	<b>Gullfaxi<sub>BASE</sub> (RRN)</b>	<b>R/C/L</b>
<b>Intro</b>	formaður sjúkraliðafélags Íslands segir að álagið á starfsfólk á heilbrigðisstofnun Íslands á síðasta ári hafi numið allt að sex milljörðum króna.	1/0/1
	<b>Gullfaxi<sub>BASE</sub> (CNN/DailyMail)</b>	<b>R/C/L</b>
<b>Intro</b>	sandra b. franks. segir að hún hafi verið í sóttkví á heimili sínu í norðurhluta lundúna. hún segir að álagið hafi komið í veg fyrir að fólk kæmist inn á heimili sín.	0/0/0
	<b>Gullfaxi<sub>BASE</sub> (CNN/DailyMail + RRN)</b>	<b>R/C/L</b>
<b>Intro</b>	það er sífellt erfiðara fyrir fólk sem búist við að fari á legudeildir á heimilum sínum og á heimili sínu. þetta segir formaður félags sjúkraliðs.	0/0/0
	<b>mT5<sub>BASE</sub> (RRN)</b>	<b>R/C/L</b>
<b>Intro</b>	formaður sjúkraliðafélags Íslands segir að faraldurinn sé loksins að sjá fyrir endann á faraldrinum. þetta segir formaður sjúkraliðafélags Íslands.	0/0/0

---

**RRN Document**

---

<b>Title</b>	LGBT-frí svæði bönnuð í Póllandi
<b>Intro</b>	Samtök sem berjast fyrir málefnum hinsegin fólks í Póllandi fagna niðurstöðu hæstaréttar stjórn-sýslulaga. Hann ógilti reglur fjögurra sveitarfélaga sem komu upp svokölluðum svæðum án hinsegin fólks.
<b>Main</b>	Hæstiréttur stjórn-sýslulaga í Póllandi hefur staðfest ógildingu svokallaðra svæða án hinsegin fólks í fjórum sveitarfélögum. Baráttusamtök fyrir málefnum hinsegin fólks fagna sigri mannréttinda í Póllandi. Herferð gegn hómófóbíu eru baráttusamtök hinsegin fólks í Póllandi, fögnuðu stórsigri í gær. Justyna Nakielska, talsmaður samtakanna segir í samtali við Fréttastofu að svæðin eru þekkt sem "svæði án hinsegin fólks (LGBT-free zones)". Justyna segir að tugir sveitarfélaga hafi samþykkt reglugerðina og hafi þáverandi formaður Mannréttindaskrifstofu Póllands kært 9 sveitarfélög. Í fimm tilvikum ákváðu lægri dómstig að þessi svæði væru ólögleg. Fjórum málum var áfrýjað til Hæstarét-tar stjórn-sýslulaga sem staðfesti að svæðin væru ólögleg í öllum fjórum tilvikum. Justyna vonar að núverandi formaður Mannréttindaskrifstofunnar haldi áfram að vinna í að ógilda reglugerðina í öllum hinum sveitarfélögunum. Reglugerðir sveitarfélaganna sneru að því að banna svokallaða upphafningu samkynhneigðar, sérstaklega í skólum. Framkvæmdarstjórn ESB lýsti því yfir að reglugerðirnar gengju gegn lögum Evrópusambandsins um að ekki mætti mismuna fólki á grundvelli kynhneigðar. Framkvæmdastjórn Evrópusambandsins hefur innleitt í samstarfssamninginn við Pólland ákvæði sem bannar fjárhagslegan stuðning við sveitarstjórnir sem hafa samþykkt svæði án hinsegin fólks". Ákvæðið í samstarfssamningnum mun án efa leiða til þess að ályktanir sveitarstjórna verða felldar úr gildi, vegna þess að fjárveitingar sambandsins nema rúmlega sjötíu tveimur milljörðum evra. Herferð gegn hómófóbíu segir að niðurstaða Hæstaréttar stjórn-sýslulaga og nýja ákvæðið í lögum ESB sé stórsigur í baráttunni fyrir lýðræði og mannréttindum.
<b>Summary</b>	Samtök sem berjast fyrir málefnum hinsegin fólks í Póllandi fagna niðurstöðu hæstaréttar stjórn-sýslulaga. Hann ógilti reglur fjögurra sveitarfélaga sem komu upp svokölluðum svæðum án hinsegin fólks.

<b>Gullfaxi<sub>BASE</sub> (RRN)</b>		<b>R/C/L</b>
<b>Intro</b>	dómstóll í póllandi hafnaði í gær kröfu framkvæmdastjórnar evrópusam-bandsins um að skerða réttindi hinsegin fólks.	1/0/1
<b>Gullfaxi<sub>BASE</sub> (CNN/DailyMail)</b>		<b>R/C/L</b>
<b>Intro</b>	mannréttindasamtök fögnuðu sigri á réttindum hinsegin fólks í póllandi í gær	1/1/1
<b>Gullfaxi<sub>BASE</sub> (CNN/DailyMail + RRN)</b>		<b>R/C/L</b>
<b>Intro</b>	mannréttindasamtök gegn mismunun í garð hinsegin fólks fagna áfanganum. mannréttindasamtök hafa felld úr gildi úrskurð hæstaréttar sem kveðinn var upp í gær.	1/0/1
<b>mT5<sub>BASE</sub> (RRN)</b>		<b>R/C/L</b>
<b>Intro</b>	formaður mannréttindaskrifstofu póllands hefur staðfest ógildingu svokallaðra svæða án hinsegin fólks. baráttusamtök fyrir málefnum hinsegin fólks fagna sigri mannréttinda í póllandi.	1/1/1

## D. Human Evaluation Examples

RRN Document		
<b>Title</b>	Algjör orkuskipti raunsær kostur	
<b>Intro</b>	Framkvæmdastjóri Vistorku á Akureyri segir raunsæjan kost að Ísland geti orðið algjörlega óháð olíu á næstu árum og þar með sjálfbært um alla orkuframleiðslu. Heildræna stefnu vanti þó í málaflokknum.	
<b>Main</b>	Fyrirtækið Vistorka hefur það að markmiði að stuðla að framleiðslu umhverfisvæns eldsneytis með sjálfbærri nýtingu hráefnis. Guðmundur Haukur Sigurðarson, framkvæmdastjóri fyrirtækisins segir tímaspursmál hvenær hætt verði að nota olíu sem orkugjafa. Sá dagur mun renna upp að við hættum að nota olíu, hvort sem það verður af okkar eigin ákvörðunum eða hún tæmist. Við Íslendingar í gegnum tíðina höfum verið ofsalega háð olíu. Málið er að við erum með bullandi framleiðslu á öðru eldsneyti. Guðmundur nefnir þar sem dæmi metan, lífdísil og vetni. Hann telur því að hægt væri að flýta banni á innflutningi á bensín- og dísilbílum til 2025. Tæknilegar áskoranir séu þó vissulega einhverjar. En að það séu einhverjar hindranir eða eitthvað ómögulegt, það er ekki. Hér á Akureyri erum við búin að framleiða sem samsvarar milljón lítrum af olíu af metani og milljón lítrum af lífdísli sem samsvarar tveimur milljón lítra af olíu erum við búin að framleiða hér á Akureyri bara úr ruslinu okkar. Lausnin í umhverfismálum sé þó ekki, að mati Guðmundar, að skipta út bílum sem nota olíu fyrir aðra sem nota umhverfisvæna orkugjafa. Það þurfi einfaldlega að fækka bílum, enda séu þeir mun fleiri á Íslandi en í öðrum Evrópuríkjum. Við þurfum einhvern veginn að vinna í þessu heildarþrógrammi í breyttum ferðavenjum. Vissulega koma rafbílur í staðinn en 230.000 rafbílur eða dísilbílur það er jafnmikið magn af bílum.	
<b>Summary</b>		
	<b>Gullfaxi<sub>BASE</sub> (RRN)</b>	<b>R/C/L</b>
<b>Intro</b>	guðmundur haukur tryggvason, framkvæmdastjóri viðskiptaþróunar hjá orku náttúrunnar, segir að stefnt sé að því að hefja framleiðslu á endurnýjanlegum orkugjöfum.	1/0/1
	<b>Gullfaxi<sub>BASE</sub> (CNN/DailyMail)</b>	<b>R/C/L</b>
<b>Intro</b>	framkvæmdastjóri segir að það er engin hindranir til að flýta banni við innflutningi á olíu.	1/1/0
	<b>Gullfaxi<sub>BASE</sub> (CNN/DailyMail + RRN)</b>	<b>R/C/L</b>
<b>Intro</b>	framkvæmdastjóri eldsneytis á akureyri segir að það þurfi að fækka bílum til að koma í veg fyrir að draga úr olíunotkun á næstu árum.	1/0/0
	<b>mT5<sub>BASE</sub> (RRN)</b>	<b>R/C/L</b>
<b>Intro</b>	forstjóri fyrirtækisins segir tímaspursmál hvenær hætt verði að nota olíu sem orkugjafa. framkvæmdastjóri fyrirtækisins segir tímaspursmál hvenær hætt verði að nota olíu.	1/1/0

---

**RRN Document**

---

<b>Title</b>	Afbókanir vegna gosloka	
<b>Intro</b>	Fyrirtækjum í ferðaþjónustu hafa nú þegar borist afbókanir vegna mögulegra gosloka. Engin virkni hefur mælst í gígnum í Meradölum frá því á sunnudag.	
<b>Main</b>	Litakóði Veðurstofu Íslands var færður úr appelsínugulum yfir í gulan í gær og að sögn Sigríðar Magneu Óskarsdóttur, náttúrvársérfræðings er gosrásin sennilega stífluð. Lítið sem ekkert sé að fréttu af gosinu. Við fylgjumst með órönum og jarðskjálftavirkninni hún er ekki mikil. Smáskjálftavirkni áfram á Reykjanesi. Óróinn hefur haldist lítill áfram þarna við gosstöðvarnar. Mikill áhugi var á Íslandi meðal erlendra ferðamanna þegar gosið hófst í byrjun mánaðar og bókanir hjá Icelandair og Play tóku meðal annars kipp. Uppbókað var í þyrluflug yfir gosstöðvarnar nær samdægurs. Birgir Ómar Haraldson framkvæmdastjóri Norðurflugs segir að töluvert hafi verið um afbókanir í gær. Það fylgir þessu en samt sem áður er það þannig að það eru mjög margir sem vilja samt sem áður sjá nýjar eldstöðvar. Þannig það hefur komið okkur aðeins á óvart hvað það er ekki eins mikið afbókað og við töldum. Birgir segir að fyrirtækið hafi látið viðskiptavinum vita af stöðunni á gosinu í gær og fólki gefið kost á að afbóka flugið. Fimmtíu prósent þeirra sem áttu bókað flug hafi valið að afbóka flugið. Ég myndi svona giska á það sé tæplega helmingur sem vill halda áfram. Þrátt fyrir að gosinu sé að öllum líkindum lokið hafa ekki margar afbókanir borist flugfélögunum síðan í gær. Ásdís Ýr Pétursdóttir, upplýsingafulltrúi Icelandair, segir að félagið hafi ekki orðið vart við afbókanir. Eftirspurn eftir ferðum til Íslands hafi verið mjög mikil síðustu misseri og útlit sé fyrir að svo verði áfram. Nadine Guðrún Yhagi, upplýsingafulltrúi Play, tekur í sama streng.	
<b>Summary</b>	Engin virkni hefur mælst á gosstöðvunum frá því á sunnudagsmorgun. Margir ferðamenn eru hættir við ferðir sínar.	
<b>Gullfaxi<sub>BASE</sub> (RRN)</b>		<b>R/C/L</b>
<b>Intro</b>	útlit er fyrir að gosinu í gosstöðvunum í gær verði lokið í næstu viku. útlit er fyrir að gosið í eyjafjallajökli hefjist í næstu viku.	1/0/0
<b>Gullfaxi<sub>BASE</sub> (CNN/DailyMail)</b>		<b>R/C/L</b>
<b>Intro</b>	í gær var greint frá því að gosinu hafi verið aflýst en í gær var greint frá því að gosinu hafi ekki verið aflýst	0/0/0
<b>Gullfaxi<sub>BASE</sub> (CNN/DailyMail + RRN)</b>		<b>R/C/L</b>
<b>Intro</b>	mikil úrkoma hefur verið á gosstöðvunum á Íslandi síðustu daga. bókað hefur verið í rúman mánuð og bókanir í gær hafa ekki verið bókaðar.	1/0/0
<b>mT5<sub>BASE</sub> (RRN)</b>		<b>R/C/L</b>
<b>Intro</b>	icelandair og play tóku meðal annars kipp. uppbókað var í þyrluflug yfir gosstöðvarnar í gær.	1/0/0



## D. Human Evaluation Examples

RRN Document		
<b>Title</b>	Íþróttir	
<b>Intro</b>	Ísland og Portúgal mætast í kvöld í undankeppni Evrópumóts karla í handbolta 2022. Þetta er fyrsti leikurinn af þremur sem liðin leika í þremur löndum á næstu átta dögum.	
<b>Main</b>	Liðin mætast í Portúgal í kvöld og er það fyrri leikur liðanna í undankeppni Evrópumótsins 2022. Svo mætast liðin öðru sinni í undankeppninni á Ásvöllum í Hafnarfirði á sunnudag og svo mætast þau líka í fyrsta leik HM í Egyptalandi á fimmtudaginn í næstu viku. Íslenska liðið er án Arons Pálmarssonar í þessum verkefnum en Guðmundur Guðmundsson, landsliðsþjálfari, segir það opna dyr fyrir aðra leikmenn. Portúgal hefur verið mjög vaxandi lið á undanförunum árum. Liðið varð í sjötta sæti Evrópumótsins í fyrra og flestir leikmenn liðsins leika saman hjá Porto í heimalandinu. Liðin mættust síðast á EM í fyrra og þá hafði Ísland betur. Leikur Íslands og Portúgal hefst klukkan 19:30 og verður sýndur beint á RÚV. Útsending hefst klukkan 19:15. Undirbúningur sænska landsliðsins fyrir HM er í uppnámi eftir að leikmaður liðsins greindist með kórónaveiruna. Liðið átti að leika gegn Svartfjallalandi í undankeppni EM í kvöld en þeim leik var frestað þar sem allt sænska liðið er komið í sóttkví. Svíar eiga að mæta Svartfjallalandi í Svíþjóð um helgina og halda svo til Egyptalands á HM eftir helgi en þær áætlanir eru allar upp í loft sem stendur. Tottenham komst í gærkvöldi í úrslit enska deildar bikarsins í fótbolta. Tottenham mætti b-deildarliði Brentford í Lundúnum og vann með tveimur mörkum gegn engu. Í úrslitum mætir Tottenham annað hvort Manchester United eða Manchester City. Þau eigast við í grannaslag í hinum undanúrslitaleiknum í kvöld en úrslitaleikurinn sjálfur fer svo fram á Wembley-leikvanginum í apríllök.	
<b>Summary</b>	Ísland og Portúgal mætast í kvöld í undankeppni Evrópumóts karla í handbolta 2022.	
<b>Gullfaxi<sub>BASE</sub> (RRN)</b>		
	<b>R/C/L</b>	
<b>Intro</b>	íslenska karlalandsliðið í fótbolta mætir portúgal í undankeppni hm í kvöld. liðið mætir portúgal í fyrsta leik í undankeppni hm í kvöld.	1/0/0
<b>Gullfaxi<sub>BASE</sub> (CNN/DailyMail)</b>		
	<b>R/C/L</b>	
<b>Intro</b>	portúgal mætir í fyrsta sinn í undankeppni evrópumótsins í knattspyrnu í kvöld. lið em mætir í fyrsta sinn í undankeppni em. lið em mætir í umspili um sæti í umspili um sæti í umspili um sæti í umspili um sæti í umspili um sæti í undankeppni em.	1/0/0
<b>Gullfaxi<sub>BASE</sub> (CNN/DailyMail + RRN)</b>		
	<b>R/C/L</b>	
<b>Intro</b>	íslenska karlalandsliðið í fótbolta mætir portúgal í umspili um laust sæti á em í umspili um laust sæti á em í egyptalandi á næstu leiktíð.	1/0/0
<b>mT5<sub>BASE</sub> (RRN)</b>		
	<b>R/C/L</b>	
<b>Intro</b>	íslenska karlalandsliðið í handbolta mætir portúgal í kvöld. ísland mætir portúgal í fyrsta leik hm í egyptalandi í næstu viku.	1/1/1