EROS:Entity-Driven Controlled Policy Document Summarization

Anonymous ACL submission

Abstract

Privacy policy documents have a crucial role in informing individuals about the collection, usage and protection of user's personal data by organizations. Policy documents are notorious for their complex and convoluted language, posing significant challenges to users who attempt to comprehend their content. In this paper, we propose an innovative approach to enhance the interpretability and readability of policy documents by using controlled abstractive summarization by enforcing critical entities using reinforcement learning.
Due to legal jargon, lengthy sentences, and intricate syntactic structures of privacy policy documents our approach first identifies critical information necessary to the user using span-based entity extraction model (`EEPD`, and use these entities to enhance the summary of the document. Our model `EROS` uses proximal policy optimization (PPO) to control the information and symantic structure of the generated summary. Our model shows massive improvement over base summarization techniques.

1 Introduction

In today's Internet era, access to information has never been so convenient. Everyday, an overwhelming amount of users are exploring the Internet horizon for entertainment or business purposes. Realizing an excellent opportunity to increase their customer base, many organizations offer their products or services in a much convenient online settings. In majority of the cases, customers require to signup to acquire the services on offer and while doing so, they have to agree to the term-and-conditions (T&C) or policy documents of the service providers.

A privacy policy is a crucial component of any organization that allows it to legally collect, process, store, and/or distribute personal information. It outlines how an organization will handle personal data and how it will comply with applicable data protection laws and regulations. Little that they know, on many occasions, customers are, advertenly or inadvertenly, granting full access to their sensitive and private data (e.g., name, contact information, location, etc.) to the service providers without reading or understanding the privacy policy document. Moreover, some companies collect data with distributional rights as well and make fortune by selling user's data to third party without their realization but with their inadvertent consent.[1,2]. The primary reason for such ignorance on the user part is their busy and packed schedule as well as lengthy and technical/legal language, which are usually difficulty to comprehend by a common user.

Motivation and Problem Definition: Privacy policies are essential for both businesses and individuals. For businesses, having a privacy policy can protect them from legal issues related to data privacy and usage. On the other hand, it provides transparency to individuals about how their personal information will be managed and protected by the organizations; thus enabling them to make informed decisions prior to registering for the service. Despite its importance, very few users read these lengthy and non-trivial documents and fall prey to their inadvertent consent.

Summarizing these documents is a straightforward remedy of the lengthy document but it needs to ensure that every aspect of the data usage/management must also be present in the summary to make it useful. However, given the complicated nature of the policy document, it's non-trivial to obtain every critical privacy-related information in a summary. As an instance, some policy docu-

[1] Brave Browser Under Fire For Alleged Sale Of Copyrighted Data

[2] Twitter fined 150m in US for selling user's data

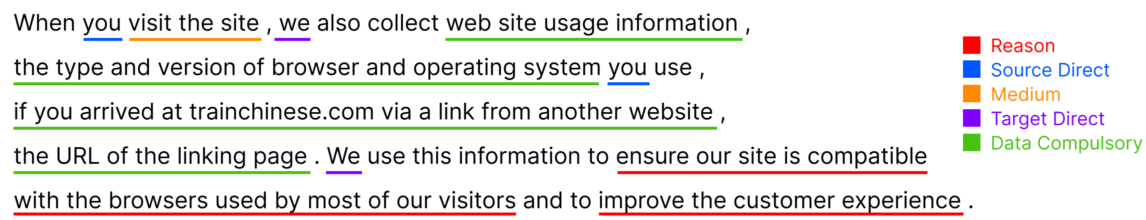

Figure 1: Example of an annotated paragraph, with the entities labelled. Here "we" belong to the entity who is directly collecting the data, therefore **target direct** and the data is being collected from "you", therefore **source direct**. When you visit the site, which acts as a medium, website usage information and the version of the browser etc, (**data**) is collected. And the **reason** for collecting data is to ensure website compatibility and improve customer experience.

ments define different data items (*viz.* name, age, contact details, etc.) at the beginning of the document but refrain themselves in reporting the management of data items until the end of the document or in different paragraph or context; thus making it challenging for any summarization system to deal with such cases. Controlled abstractive summarization techniques He et al. (2020), Liu and Chen (2021), Zhang et al. (2023) can potentially enhance accessibility and transparency by generating concise and coherent summaries of policy documents, which can make the content more comprehensible and manageable for the general public. Acknowledging the severity of the problem, in this paper, we propose an **Entity-dRiven cOntrlled policy document Summarization** system *aka.* **EROS**. EROS operates in two stages: 1) it extracts various entities or data items and their rationales through a BERT and XLNet-based entity-extraction module; and 2) leveraging the extracted entities, it mandates a BART-based summarization module to include these entities and their rationales through a proximal policy optimization (PPO) framework.

We develop a dataset, namely PD-Sum, of 1900 policy documents and manually annotated them with abstractive summaries along with privacy-related entities and their rationales. At first, we mark all entities present in the document and also identify what, why, and how they are being collected. To achieve this, we proposed and followed a schema for the identification of critical privacy-related information in a policy document as depicted in Figure 2 (c.f. Section 3 for details). It includes the ***data*** being collected, who would be the ***source*** of data, through which ***medium*** data will be collected, who will consume (***target***) the data, and what is the ***reason*** of data collection. An example with annotated entities is shown in Figure 1. In the next step, we write a summary of the document mandating the presence of the entities and their rationales along with other relevant information.

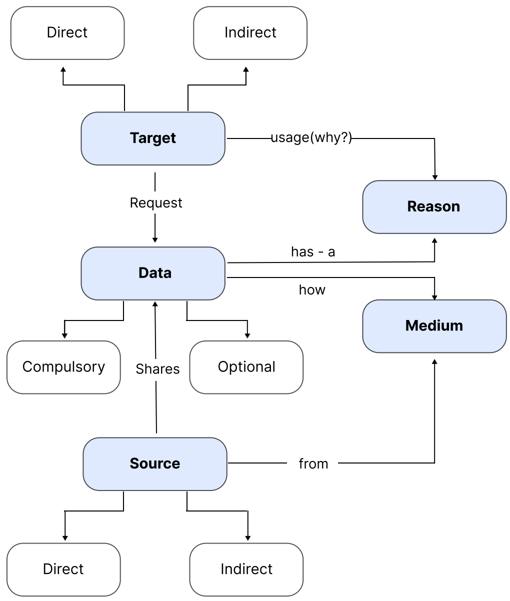

Figure 2: Relation between annotation labels

Our experimental results demonstrate that EROS achieves state-of-the-art performance on the proposed PD-Sum dataset against several baseline systems. We also perform qualitative and error analyses to assess the capability of EROS in ensuring various aspects of the privacy-related information in the generated summaries.

Contribution: The main contributions of this paper are summarized as follows:

- We propose a BART-based entity-driven controlled policy document summarization (EROS) to mitigate the concerns of general public over the data privacy and security issues.
- To identify privacy-related relevant information in a policy document, we developed an entity

extraction model, **E**ntity **E**xtraction from **P**olicy **D**ocuments (EEPD).

- We introduce a personalized loss function and a reinforcement learning framework using Proximal Policy Optimization (PPO) to managed the relevance and length of the generated summaries.
- We introduce a new dataset (PD-Sum) of privacy policy documents with their summaries and privacy-bounded entities and rationales.
- We also establish performance benchmarks for the proposed approach against several baselines.
- Finally, we perform qualitative and error analyses to assess the quality of summaries.

Reproducibility: Code and datasets will be released on acceptance.

2 Related Work

Pretrained encoders have become pivotal in recent summarization approaches. Liu et al. (2020) introduced a BERT-based unsupervised text summarization model, achieving state-of-the-art performance on benchmark datasets. Wang et al. (2021) explored fine-tuning strategies for language models like BERT and GPT, revealing substantial performance gains through limited labeled data utilization. Dong et al. (2021) proposed a hierarchical transformer model for summarizing lengthy documents, achieving leading results on multiple benchmarks. Zhang et al. (2020) devised PEGASUS, leveraging gap sentence extraction and transformer-based gap filling pre-training, attaining state-of-the-art performance on various benchmarks. These studies depict the impactful role of pretrained encoders in advancing summarization techniques.

Entity extraction is a fundamental task in information extraction. The problem has been modelled in multiple ways such as sequence labelling Francis et al. (2019); Lin et al. (2020); Bui et al. (2021), span level prediction Eberts and Ulges (2020); Zhong and Chen (2020); Zhu and Li (2022), question answering Li et al. (2020) as well as dependency parsing task Yu et al. (2020).

Recently, a paradigm shift has been observed from sequence labelling task to span-based prediction of entities. In span-based task, such as Eberts and Ulges (2020), all possible spans are selected and further classified whether that span represents an entity or not followed by relation classification, if required. To tackle the problem of exact spans being treated correctly and partial spans being treated incorrectly, Zhu and Li (2022), proposes a way to regularize span-based prediction tasks. The annotated spans are assigned full probability and the nearby tokens are also assigned some probability of being correct. Zhong and Chen (2020) extracts entities along with relation instead of the traditional approach of extracting entities and then using the extracted entities for relation classification.

Reinforcement learning (RL) has gained traction in summarization. Approaches combine supervised learning and RL for abstractive summarization, while hierarchical RL merges word and sentence operations with saliency-based attention Gu and et al. (2016); Paulus and et al. (2017); Wang and et al. (2018); Wan and et al. (2018). Rondeau and et al. (2018) introduced RL-driven translation with simulated human feedback. Liu and et al. (2020) addressed RL's reward scarcity using human feedback. Gunasekara and et al. (2021) presented a versatile framework using RL for abstractive summarization through question-answering rewards. These efforts highlight RL's effectiveness in improving summarization.

In comparison to existing works, our paper introduces a novel dimension to the field of summarization. While various studies have focused on leveraging pretrained encoders, reinforcement learning, and modified loss functions, our approach integrates these elements. Our model incorporates a penalty mechanism in addition to a refined BART architecture for controlled summarization. In order to provide more exact and controlled summaries, this combined technique builds on the advantages of each component resulting in more precise and controlled summary generation. Additionally, our method uniquely incorporates insights from an Entity Extraction task, enhancing the model's ability to capture information from policy documents.

3 Dataset Construction

To the best our knowledge, the domain of privacy-driven policy document summarization has not been studied so far; hence, we recognize the need of a dataset to facilitate our research and thereby, develop PD-Sum, tailor-made dataset for the policy document summarization.

Data collection and Filtering: We collect policiy documents of different websites curated by Amos and et al. (2021). These documents outline how websites manage (i.e., collect, use, or disclose) personal information. After collection, we observe several issues and hence, apply a filter to discard policy document as follows:

- Many websites have identical policy documents, we discard all but one.
- In case there are URLs linking to other websites, disregard them.
- Skip documents that lack meaningful/significant information or are incomplete.
- Refrain from including any policy content that is not relevant to the topic at hand.

Subsequent to the filtering process, 1920 policy document remains in `PD-Sum`.

Data Annotation: To facilitate the entity-driven controlled summarization, we need two sets of annotations: a) identification of privacy-related entities; and b) a summary of the document. Considering the users' concern, we identified five fundamental entities regarding the data privacy and security and proposed a schema (c.f. Figure 2) to capture their relationships:

- **Data:** It defines the type of information that an organization usually collects – *name*, *email*, *contact number*, *address*, *location*, *photos*, *system details*, *browsing history*, *search queries/patterns*, *keystrokes*, etc. Further, we observe that some of these data are compulsory as part of the service agreement, while others are optional and user can deny the access without any interruption in service.
- **Source:** It signifies the provider of the information. While majority of the time, the user (e.g., '*you*') is the direct source, in some cases, source can be indirect, e.g., "inviting your friends to join the website by sharing contact information "(*friends* will be source indirect), "requiring someone else information to ship products to their address"(*someone else* is source indirect). These entities are very low in number and majorly associated with sharing someone else information.
- **Medium:** It defines the way data is collected such as '*while visiting the website*', '*responding to a survey*', '*filling a form*', etc.
- **Target:** It specifies who will consume the data. Similar to the source entity, a target can be direct (*the organziation itself*) or indirect (*any third-party vendor outside the organization*). Though the direct target is somewhat benign as the users know their data are being used for some specific purposes by the service-provider, the indirect target can be extremely detrimental as there is no transparency about the usage of data in an unknown capacity.
- **Reason:** It clarifies the purpose of data collection by the parent organization such as '*improving customer experience*'. We observe that with indirect targets, reasons are usually hidden or extremely vague.

Following the above schema, two annotators[3] with good English proficiency annotated the whole dataset using LabelBox[4] as the annotation tool. At first, we tokenize the sentences using NLTK tokenizer (Bird et al., 2009), and subsequently, for each identified entity, we record their start and end indices as span. Further, to ensure the consistency of annotations between them, the annotators independently annotated a small set of documents in the pilot phase and discussed their common understanding. Next, they annotated 10 documents separately and achieved a Cohen Kappa inter-annotator agreement score of 0.74. Subsequently, we annotate the complete dataset of 1920 documents. In total, we annotate 8000 sentences with 9094 distinct entity labels. An illustrative example of annotation for a paragraph of the document is presented in Figure 1.

In the second stage, we annotate each document with a concise summary. We ask annotators to follow a simple guideline of inducing relevant data-related phrases in their summary while maintaining the gist of the document as concise and crisp as possible. We further instruct them to write summaries is simpler and regular words and avoid the usage of fancy and sophisticated words as much as possible. On average, policy documents have 1700 tokens, while the annotated summaries contain 203 tokens. Examples of annotated summaries are furnished in Appendix.

4 Proposed Methodology - `EROS`

Our proposed controlled summarization model, `EROS`, is depicted in Figure 3. It works in two stages. At first, we train an entity extraction model that aims to predicts all spans of entities in a given document. This model employ BERT-based span prediction framework with the contrastive loss. Additionally, we supplement the prediction via an entity classification model in a joint learning setup. In the second stage, we employ a BART-based sum-

[3] Annotators were undergraduate student volunteers and in the age group of 20-30.

[4] https://labelbox.com/

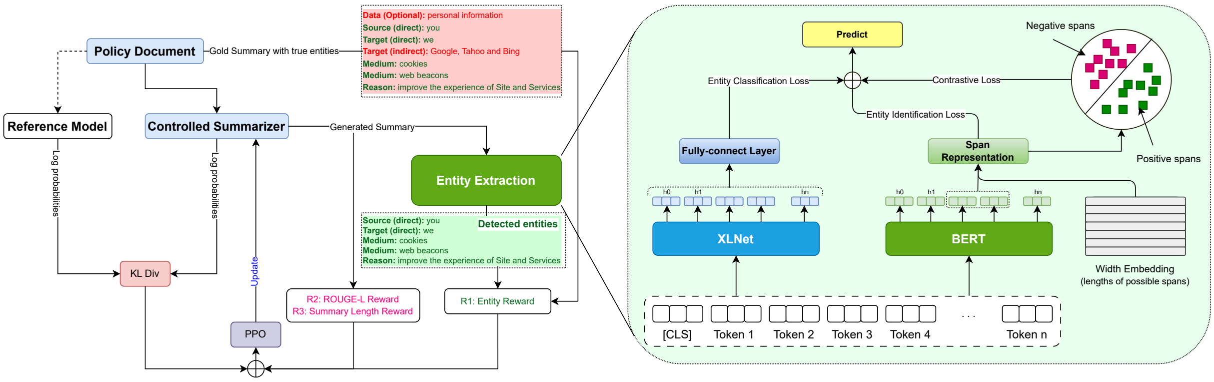

Figure 3: **Left:** Proposed Model for EROS. The reference model is a frozen pre-trained BART-based model with modified loss. We initialize the controlled summarization model in a similar way, which is subsequently updated through a PPO framework on a combination of rewards and KL-divergence loss. **Right:** Entity extraction model jointly learns a entity classification and entity identification module with the assistance of contrastive loss. Further to minimize the effect of false positives in identification, we supplement it with a entity classification module in a joint framework.

marizer model to generate a summary. To assess the quality of generated summary and to ensure induction of critical entity information in the summary, we introduce our pretrained entity extraction module in the pipeline. The extracted entities are then compared with the true (gold) entities of the reference summary and an entity reward is computed. Moreover, we compute two other rewards (ROUGE-L and summary length) to ensure a relevant and concise summary. The accumulated reward is then added with the KL-divergence score between the log-probabilities of the controlled summarizer and a reference model (a BART-based summarizer trained on a huge out-of-domain corpus). Subsequently, we employ PPO to update the controlled summarizer. This process repeats for a few step till (near-)convergence. In the following subsections, we elaborate on these steps.

4.1 Entity extraction module

Recent years have seen a paradigm shift in the task of entity recognition from token-level tagging –which conceptualizes it as a sequence labelling task– to span-level prediction (Fu et al., 2021).

A span in a sentence is represented by the start and end token of a sentence. Given a sentence $X = \{x_1, \cdots, x_n\}$ with n tokens, we define a span of an entity as $s_i^y = \{x_{b_i}, x_{b_i+1}, \cdots, x_{e_i}\}$, where b_i and e_i denote the start and end index of the span s_i with a corresponding tag $y \in \{$*source-direct, source-indirect, data-optional, data-compulsory, medium, target-direct, target-indirect, reason*$\}$.

Spans come in different lengths. To avoid overfitting for a particular length, we adopt an enumeration strategy, where all the possible m spans with a maximum length l are being considered as valid spans for predicting entities. For example, in the sentence, "*we will collect name*" with maximum span length as 4, possible spans are: $s_1^y = \{x_1, x1\}$, $s_2^y = \{x_1, x2\}$, $s_3^y = \{x_1, x3\}$, $s_4^y = \{x_1, x4\}$, $s_5^y = \{x_2, x2\}$, $s_6^y = \{x_2, x3\}$, $s_7^y = \{x_2, x4\}$, $s_8^y = \{x_3, x3\}$, $s_9^y = \{x_3, x4\}$, and $s_{10}^y = \{x_4, x4\}$. From gold labels, we know that out of these 10 spans, only s_1 and s_4 are valid spans; therefore, their y labels will be source-target and data-compulsory, respectively. For all other spans, the y labels will be "invalid (0)".

We obtain token representation from BERT and subsequently, compute span embeddings as $\mathbf{z}_i^b = [\mathbf{h}_{b_i}; \mathbf{h}_{e_i}]$ Additionally, to provide information regarding the width of each span, we induce a learnable width encoding vector z_i^w according to their width. Thus, the final representation becomes $\mathbf{s}_i = \left[\mathbf{z}_i^b; \mathbf{z}_i^w\right]$.

Our initial experiments showed encouraging results; however, we also observe a significant number of false-positive, especially, for a sentence with no entity at all, in our predictions. To mitigate such issue, we incorporate a binary classification task to identify if the sentence contains an entity in a joint learning framework.

Contrastive Loss: We compute a similarity score between pairs of input spans, and then minimize the distance between similar pairs while maximizing the distance between dissimilar pairs using contrastive loss.

$$\mathrm{P}\left(y \mid s_i\right) = \frac{\operatorname{score}(\mathbf{s}_i, \mathbf{y})}{\sum_{y' \in \mathcal{Y}} \operatorname{score}(\mathbf{s}_i, \mathbf{y}')}$$

where score(·) is a function that measures the compatibility between a specified label and a span:

$$\text{score}(\mathbf{s}_i, \mathbf{y}_k) = \exp\left(\mathbf{s}_i^T \mathbf{y}_k\right)$$

where $\mathbf{s}_i$ denotes the span representation and $\mathbf{y}_k$ is a learnable representation of the class k.

Span Prediction

Finally, the span representations s_i are fed into a softmax function to get the probability considering the label y. For optimization, we combine the three losses – entity identification cross-entropy loss (ℓ_1), binary classification loss (ℓ_2), and the contrastive loss (ℓ_3) through the following weighting mechanism.

$$\mathcal{L}^e = \alpha_1 \ell_1 + \alpha_2 \ell_2 + \alpha_3 \ell_3, \quad \text{where} \sum \alpha_i = 1$$

4.2 Entity-Driven Controlled Summarization

As the foundational model, we employ BART in our experiment. Further, we induced a modified loss function especially designed for the entity-driven controlled summary generation. To elaborate, we obtain gold entities (e_i) from the PD-Sum dataset and integrated them into the loss function of the BART model. This entails augmenting the traditional cross-entropy loss with a penalty component derived from the extracted entities. It enables the BART model to comprehend the presence and importance of entities in the summaries, thereby refining its summary generation capabilities while maintaining control over the process. Mathematically it can be seen as follows:

$$\begin{aligned} \text{CE} &= -\sum (y \cdot \log(x)) \\ TP &= \sum_{e_i} (1.0 - \text{step}(e_i \in S_G)) \\ \mathcal{L}^s &= \lambda \cdot \text{CE} + (1 - \lambda) \cdot \text{TP} \end{aligned}$$

where CE, TP, λ, and S_G are the cross-entropy, token penalty score, weight of the loss, and the generated summary, respectively. We compute TP by penalizing the model for each missing entity e_i in S_G. The step function will return 1 only if the entity is part of the summary, else, a value of 0 will be returned.

Further to supplement the controlled summary generation process, we adopted a feedback mechanism, in the form of reinforcement learning, to reward/penalize the model for inducing/not-inducing the privacy-related entities in the summaries. We use proximal policy optimization (PPO) to enforce the model to improve the generation quality. First introduced by Schulman et al. (2017), PPO refines policy adjustments by combining ratio-based enhancement with a clipped surrogate objective; thus, ensuring controlled updates. Incorporating an auxiliary value function, PPO enhances policy updates by estimating advantages and rewards more accurately, particularly in complex scenarios.

The proposed model shown in Figure 3, contains a policy model (i.e., the controlled summarizer model that is being trained), a reference model, a reward model, and a value function. The value function is used to describe the reward at timestep t. On the other hand, the reference model is used to calculate the KL-divergence between the original model and the policy model. The main idea is to ensure that the active model does not deviate a lot from its original distribution.

Reward calculation: We compute three rewards to maintain the coverage, conciseness, and relevance in the summary. The coverage reward ensures the readability of the generated summary –we compute ROUGE-L score, which is based on longest common subsequence (LCS) between two sequences, as the first reward ($R1 =$ ROUGE-L(S_G, S_R)). A longer LCS indicates that generated summary conveys similar meaning and concepts as the reference summary. The conciseness reward ($R2$) limits the model to generate adequate lenght summary and avoid generating lengthy jargons. The following equation defines $R2$:

$$R_2 = \frac{1 - |(len(S_G) - len(S_R))|}{max(len(S_G), len(S_R))}$$

Finally, we compute the entity reward ($R3$) as follows: Let E_{total} be the total number of entities predicted from the generated summary, and $E_{correct}$ and $E_{incorrect}$ be the number of entities present and not present in the gold summary respectively.

$$R_3 = \frac{E_{correct} - \beta * E_{incorrect}}{E_{total}}$$

where β is a negative factor for penalizing incorrect entities. We empirically set $\beta = 0.3$ for our experiments.

5 Experiments and Results

We train EROS on 1536 documents, while we use 385 documents for evaluating the performance. We evaluate EROS and the entity extraction module separately and their results are furnished in Tables 2 and 1, respectively. We also perform extensive comparative analysis against the following baselines for both models. In all cases, we re-train/fine-tune these models on PD-Sum.

Baselines:

- **Entity Extraction:** We evaluate EEPD against six entity extraction models covering both sequence-labelling and span-based frameworks:
  - **BERT Devlin et al. (2018)**: Fine-tuning BERT on sequence labelling task.
  - **SpanBERT Joshi et al. (2020)**: Pre-training method designed to better represent spans and predict text spans. Finetuned on sequence labelling task
  - **PrivBERT Srinath et al. (2021)**: PrivBERT is a privacy policy language model. It is pre-trained on 1 million privacy policies starting with the pre-trained Roberta model. We fine-tune this on sequence labelling task.
  - **Boundary Smoothing Zhu and Li (2022)**: Model based on span extraction.
  - **PURE Zhong and Chen (2020)**: Model based on span extraction.
  - **SPERT Eberts and Ulges (2020)**: SPERT is a joint entity and relation extraction model based on BERT, which adds local context to relate the extracted spans better. Model based on span extraction
- **Entity-Driven Controlled Summarization:** For the controlled summarization model, we compare EROS's effectiveness against the following baseline approaches:
  - **Extractive Oracle Hirao et al. (2017):** Employs extractive methods to directly gather essential information from the source text for summarization. This model offers efficiency by avoiding new sentence generation, though it can miss overall context and flow.
  - **Bert2Bert Chen et al. (2022):** This model capitalises on the strengths of BERT's pre-training on a vast amount of textual data, enabling it to capture the contextual information present in the source text and generate a coherent summary.
  - **T5-Summarizer Raffel et al. (2020):** The Text-to-Text Transfer Transformer, employs a Transformer-based structure centered around the text-to-text methodology.
  - **BART-Summarizer Lewis et al. (2019):** A denoising autoencoder for pretraining sequence-to-sequence models. Trained by corrupting text using a noising function, then learning to reconstruct it.
  - **PEGASUS Zhang et al. (2020):** Pre-trained transformer-based sequence-to-sequence architecture designed by Google AI. Model uses a novel pre-training objective known as "gap-sentences generation".

		Precision	Recall	F1-score
BIO	BERT	0.31	0.38	0.34
	SpanBERT	0.31	0.39	0.35
	PrivBERT	0.37	0.44	0.40
Span Based	SPERT	0.10	**0.68**	0.17
	Boundary Smoothing	0.40	0.48	0.44
	PURE	0.35	0.12	0.17
	SpanNER	0.49	0.56	0.52
	SpanNER + Identification	0.47	0.66	0.55
	EEPD- Identification	0.48	0.57	0.52
	EEPD	**0.54**	0.62	**0.58**

Table 1: Results of entity extraction model vs baselines

Result Analysis:

Table 1 contains the comparative result of EEPD and various baselines. PrivBert reports the best F1-score of 0.40 in the sequence-labelling framework (i.e., in a BIO setup), whereas, BoundarySmoothing yields +4% better F1-score at 0.44 in the span-based setting. Further, SpanNER, with the identification module, records the best F1-score of 0.58 among all baselines. In comparison, EEPD reports the state-of-the-art performance at 0.58 F1-socre – an increment of +3% over the best baseline. We also observe the effect of the entity identification module on the overall performance – a decrement of -6% is observed on removing the identification component from EEPD.

For the controlled summarization task, we furnish the results in Table 2. We compute the traditional ROUGE, METEOR, and BLEU scores to evaluate the generated summaries. Among all baselines (except with the modified loss, TP (c.f. Section 4.2)), BART reports the best performance across the three metrics – ROUGE-L score (0.44), BLEU-4 (0.20), and METEOR (0.33). Further, we observe that the incorporation of modified TP loss obtains comparable results in majority of the cases and better several setups – ROUGE-L: T5, BART, and PEGASUS improved; BLEU-4: PEGASUS

Model	Rouge			BLEU				METEOR
	R1	R2	RL	B1	B2	B3	B4	
Extractive Oracle	0.43	0.30	0.42	0.14	0.12	0.11	0.10	0.25
Bert2Bert	0.25	0.04	0.22	0.22	0.08	0.10	0.16	0.25
T5-Summarizer	0.44	0.24	0.42	0.32	0.24	0.19	0.17	0.34
PEGASUS	0.35	0.17	0.32	0.18	0.12	0.09	0.08	0.23
BART	0.46	0.29	0.44	0.31	0.25	0.22	0.20	0.33
T5-Loss	0.45	0.26	0.43	0.32	0.24	0.19	0.17	0.34
BART-Loss	0.48	0.31	0.46	0.31	0.25	0.22	0.20	0.34
PEGASUS-Loss	0.38	0.21	0.36	0.23	0.17	0.14	0.12	0.26
EROS	**0.512**	**0.332**	**0.495**	**0.4156**	**0.329**	**0.284**	**0.254**	**0.424**

Table 2: Rouge, Bleu and Meteor scores for baselines and our proposed EROS model.

improved; and METEOR: BART and PEGASUS improved. On the other hand, EROS yields the best scores across all metrics – improvement of +3% in ROUGE-L (0.49), +5% in BLEU-4 (0.25), and +8% in METEOR (0.42).

Model	INFO	CON	FLU	EC
BART	3.0	3.35	3.75	2.90
BART-Loss	3.75	**3.70**	4.0	3.40
EROS	**4.20**	3.15	**4.05**	**4.15**

Table 3: Human Evaluation Results INFO ,CON, FLU, and EC denote Informative, Concise, Fluent, and Entity Coverage respectively.

Human Evaluation: We performed human evaluation on a subset of randomly chosen samples from the PD-Sum's test set. We compare the summaries of EROS and two baselines i.e., BART and BART-Loss. We ask our evaluators to assess the generated summaries against the reference summaries on four parameters – the informativeness of the summary (INF), its conciseness (CON), its fluency and grammatical correctness (FL), and the inclusion of relevant entities (EC). For each parameter, all evaluators assisgns a rating on a scale of 1 (worst) to 5 (best) based on the quality of the summaries. Subsequently, we aggregate the scores through averaging and report the observations in Table 3. We observe that EROS outperforms the other two baseline models into three out of four metrics. It records comparatively inferior score for conciseness, suggesting that EROS's summaries are relatively lengthier than others. However, it is better in informativeness, grammatical correctness, and inclusion of relevant entities.

6 Conclusion

In this work, we presented a novel approach for abstractive summarization of privacy policy documents. Our approach aimed to address the challenge of generating controlled and informative summaries that capture the essence of complex privacy policies. To achieve this, we introduced a customized loss function and incorporated a reinforcement learning framework, enabling us to optimize the relevance of the generated summaries. To facilitate the evaluation and advancement of research in this domain, we also introduced a new datasets for controlled summarization generation. The experimental results obtained from our comprehensive evaluations highlight the effectiveness of the proposed approach. Our model achieved state-of-the-art performance on the custom dataset. The controlled generation of summaries allows for improved accessibility and transparency for users, enabling them to quickly grasp the key points of privacy policies without getting overwhelmed by excessive information. The findings of our work demonstrate the potential of our approach to make a significant impact in the field of privacy policy summarization. By addressing the critical need for concise and user-friendly representations of privacy policies, we contribute to enhancing user understanding. The implications of our work extend to various domains where privacy policies play a crucial role, including data protection, online services, and legal compliance.

7 Limitation

Our model uses sensitive information about company privacy policies to train. It may contain personal private information regarding internal matters of a company. A reinforcement model is very computation extensive and takes long time to train.

References

Ryan Amos and et al. 2021. Privacy policies over time: Curation and analysis of a million-document dataset. In *Proceedings of the Web Conference 2021*, pages 1–12. ACM.

Steven Bird, Ewan Klein, and Edward Loper. 2009. *Natural language processing with Python: analyzing text with the natural language toolkit*. " O'Reilly Media, Inc.".

Duc Bui, Kang G Shin, Jong-Min Choi, and Junbum Shin. 2021. Automated extraction and presentation of data practices in privacy policies. *Proc. Priv. Enhancing Technol.*, 2021(2):88–110.

Cheng Chen, Yichun Yin, Lifeng Shang, Xin Jiang, Yujia Qin, Fengyu Wang, Zhi Wang, Xiao Chen, Zhiyuan Liu, and Qun Liu. 2022. bert2BERT: Towards reusable pretrained language models. In *Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 2134–2148, Dublin, Ireland. Association for Computational Linguistics.

Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. 2018. Bert: Pre-training of deep bidirectional transformers for language understanding. *arXiv preprint arXiv:1810.04805*.

Li Dong, Furu Wei, Ming Zhou, and Ke Xu. 2021. Hierarchical transformers for long document summarization. In *Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics*.

Markus Eberts and Adrian Ulges. 2020. Span-based joint entity and relation extraction with transformer pre-training.

Sumam Francis, Jordy Van Landeghem, and Marie-Francine Moens. 2019. Transfer learning for named entity recognition in financial and biomedical documents. *Information*, 10(8):248.

Jinlan Fu, Xuanjing Huang, and Pengfei Liu. 2021. SpanNER: Named entity re-/recognition as span prediction. In *Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing (Volume 1: Long Papers)*, pages 7183–7195, Online. Association for Computational Linguistics.

Jiatao Gu and et al. 2016. Deep reinforcement learning for sequence-to-sequence models. *arXiv preprint arXiv:1609.06018*.

Chamara Gunasekara and et al. 2021. Using question answering rewards to improve abstractive summarization. In *Findings of the Association for Computational Linguistics: EMNLP 2021*, pages 518–526.

Junxian He, Wojciech Kryscinski, Bryan McCann, Nazneen Rajani, and Caiming Xiong. 2020. Ctrlsum: Towards generic controllable text summarization. *ArXiv*, abs/2012.04281.

Tsutomu Hirao, Masaaki Nishino, Jun Suzuki, and Masaaki Nagata. 2017. Enumeration of extractive oracle summaries. In *Proceedings of the 15th Conference of the European Chapter of the Association for Computational Linguistics: Volume 1, Long Papers*, pages 386–396, Valencia, Spain. Association for Computational Linguistics.

Mandar Joshi, Danqi Chen, Yinhan Liu, Daniel S Weld, Luke Zettlemoyer, and Omer Levy. 2020. Spanbert: Improving pre-training by representing and predicting spans. *Transactions of the Association for Computational Linguistics*, 8:64–77.

Mike Lewis, Yinhan Liu, Naman Goyal, Marjan Ghazvininejad, Abdelrahman Mohamed, Omer Levy, Ves Stoyanov, and Luke Zettlemoyer. 2019. Bart: Denoising sequence-to-sequence pre-training for natural language generation, translation, and comprehension.

Xiaoya Li, Jingrong Feng, Yuxian Meng, Qinghong Han, Fei Wu, and Jiwei Li. 2020. A unified MRC framework for named entity recognition. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, pages 5849–5859, Online. Association for Computational Linguistics.

Bill Yuchen Lin, Dong-Ho Lee, Ming Shen, Ryan Moreno, Xiao Huang, Prashant Shiralkar, and Xiang Ren. 2020. Triggerner: Learning with entity triggers as explanations for named entity recognition. *CoRR*, abs/2004.07493.

Fei Liu and et al. 2020. Learning to summarize from human feedback. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*.

Yang Liu, Ming Li, Xin Liu, Yao Chen, and Maosong Sun. 2020. Presumm: A bert-based unsupervised text summarization model. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*.

Zhengyuan Liu and Nancy F. Chen. 2021. Controllable neural dialogue summarization with personal named entity planning. In *Conference on Empirical Methods in Natural Language Processing*.

Romain Paulus and et al. 2017. A deep reinforced model for abstractive summarization. In *Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing*.

Colin Raffel, Noam Shazeer, Adam Roberts, Katherine Lee, Sharan Narang, Michael Matena, Yanqi Zhou, Wei Li, and Peter J. Liu. 2020. Exploring the limits of transfer learning with a unified text-to-text transformer.

Mathieu-Auguste Rondeau and et al. 2018. Reinforcement learning for bandit neural machine translation with simulated human feedback. In *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing*.

John Schulman, Filip Wolski, Prafulla Dhariwal, Alec Radford, and Oleg Klimov. 2017. Proximal policy optimization algorithms. *CoRR*, abs/1707.06347.

Mukund Srinath, Shomir Wilson, and C Lee Giles. 2021. Privacy at scale: Introducing the PrivaSeer corpus of web privacy policies. In *Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing (Volume 1: Long Papers)*, pages 6829–6839, Online. Association for Computational Linguistics.

Xiaojun Wan and et al. 2018. Improving abstractive document summarization with salient information modeling. In *Proceedings of the 27th International Conference on Computational Linguistics*.

Chenguang Wang and et al. 2018. A hierarchical reinforced sequence operation method for abstractive summarization. In *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing*.

Liyuan Wang, Yue Zhang, Fuli Xu, and Xu Chen. 2021. Fine-tuning pretrained language models: Weight initializations, data orders, and early stopping. *arXiv preprint arXiv:2103.05447*.

Juntao Yu, Bernd Bohnet, and Massimo Poesio. 2020. Named entity recognition as dependency parsing. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, pages 6470–6476, Online. Association for Computational Linguistics.

Jingqing Zhang, Yao Zhao, Mohamed Saleh, and Peter Liu. 2020. Pegasus: Pre-training with extracted gap-sentences for abstractive summarization. In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing*.

Yusen Zhang, Yang Liu, Ziyi Yang, Yuwei Fang, Yulong Chen, Dragomir Radev, Chenguang Zhu, Michael Zeng, and Rui Zhang. 2023. Macsum: Controllable summarization with mixed attributes.

Zexuan Zhong and Danqi Chen. 2020. A frustratingly easy approach for entity and relation extraction. *arXiv preprint arXiv:2010.12812*.

Enwei Zhu and Jinpeng Li. 2022. Boundary smoothing for named entity recognition. In *Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 7096–7108, Dublin, Ireland. Association for Computational Linguistics.

A Appendix

A.1 Examples

The generated summaries produced by various models for comparative analysis are presented below.

Gold Summary

DSI may collect the following information when requesting product information: name and job title, contact information including email address, demographic information such as postcode, preferences and interests. The information may be used by DSI to improve their products and services. A cookie is a small file that requests permission to be placed on the hard drive of your computer. When you agree, the file is added, and the cookie analyses web traffic or notifies you when you visit a specific site. DSI employs traffic log cookies to determine which pages are being accessed. This website may contain links to other relevant websites. However, once you use these links to leave DSI, you should be aware that DSI has no control over the other website. Unless we have your permission or are required by law, DSI will not sell, distribute, or lease your personal information to third parties.

BART

Generated Summary
When requesting product information, may collect the following information: name and job title, contact information including email address, demographic information such as postcode, preferences and interests. We only use this information for statistical analysis purposes and then the data is removed from the system. Our website may contain links to other websites of interest. Once you have used these links to leave our site, you should note that we do not have any control over that other website.

BART-Loss

Generated Summary
When requesting product information, may collect the following information: name and job title, contact information including email address, demographic information such as postcode, preferences and interests. We require this information to understand your needs and provide you with a better service, and in particular for the following reasons: Internal record keeping. We may use your personal information to send you promotional information about third parties which we think you may find interesting if you tell us that you wish this to happen. If you believe that any information we are holding on you is incorrect or incomplete, please write to or email us as soon as possible. Our website may contain links to other websites of interest. However, once you have used these links to leave our site, you should note that we do not have any control over that other website.

EROS

Generated Summary
When requesting product information, may collect the following information: name and job title, contact information including email address, demographic information such as postcode, preferences and interests. We require this information to understand your needs and provide you with a better service, and in particular for the following reasons: Internal record keeping. We may use your personal information to send you promotional information about third parties which we think you may find interesting if you tell us that you wish this to happen. If you believe that any information we are holding on you is incorrect or incomplete, please write to or email us as soon as possible.

A.2 Hyperparameters

Parameters which are not declared are set to default values.

Hyper-parameter	Value
max_length	1024
batch size	4
learning_rate	5e-5
weight_decay	0.01
fp16	True

Table 4: Hyperparameter for training summarization model

Hyper-parameter	Value
max_length	1024
early_stopping	True
num_beams	3
do_sample	False
use_cache	True
length_penalty	0.3
top_k	10
top_p	0.9

Table 5: Hyperparameter for generating text from summarization model

Hyper-parameter	Value
n_class	10
BERT_DROPOUT	0.2
XLNET_DROPOUT	0.2
LR	1e-5
MAXLEN	512
MAXNORM	1.0
batchSize	4
max_spanLen	10
spanLen_emb_dim	300

Table 6: Hyperparameter for training EEPD

Hyper-parameter	Value
warmup	0.1
learning_rate	5.41e-6
adaptive_kl_coef	True
gamma	0.99
max_seq_length	1024
min_new_tokens	200
top_p	0.9

Table 7: Hyperparameter for EROS training and generation text