LLM-Based Privacy Policy Analysis and TL;DR Generation

Done By: Diya D Bhat

Problem statement

Privacy policies are often extremely long, complex, and written in legal language that most users do not read or understand. These documents may span dozens of pages and contain critical information about what types of personal data are collected, tracked, stored, and shared with third parties. As a result, users are often unaware of the privacy risks associated with using digital platforms.

The goal of this project is to design a Generative AI–based system that can generate a TL;DR (Too Long; Didn't Read) summary of large privacy policy documents (50 pages or more), specifically highlighting what kinds of user data are collected and shared, so that privacy risks become easier to understand.

Abstract

This project presents a Privacy Policy TL;DR system built using Hugging Face transformer model to summarize large privacy policy documents and assess their privacy risk. Due to the limitations of directly summarizing long legal documents, a keyword-based extraction step is introduced to isolate privacy-critical sections related to personal data collection, tracking, and third-party sharing. The extracted content is then summarized using a transformer-based abstractive summarization model.

To further enhance interpretability, a privacy risk scoring mechanism classifies documents into Low, Moderate, or High risk based on the presence of sensitive data indicators such as financial information, location tracking, device identifiers, and third-party data sharing. The system is capable of processing very large documents (including converted 100+ page policies) and was tested on real-world privacy policies such as Meta's Privacy Policy and Tata Power's Privacy Policy (screenshots of model output attached at the end).

Documentation

1. Understanding the Problem

Through experimentation, it was observed that directly summarizing full privacy policies often produces vague results that focus on company descriptions rather than concrete data collection practices. Privacy policies also contain large amounts of legal and procedural text that dilute important information. Therefore, simply applying a summarization model was insufficient to achieve meaningful results.

2. Approach and System Design

To address this, the project follows a two-stage intelligent pipeline:

Stage 1: Keyword-Based Data Extraction

Before summarization, the document is filtered to extract only privacy-critical content. This is done using a curated set of high-risk keywords, including:

- Personal identifiers (email, phone number, username)
- Financial data (credit card, billing, transactions)
- Location & tracking data (GPS, IP address, cookies)
- Device identifiers (device ID, advertising ID)
- Behavioral data (browsing history, usage patterns)
- Biometric & media data (voice, photos, videos)
- Third-party sharing & advertising

Only paragraphs containing these keywords are passed to the summarization model.
This ensures the model focuses only on what data is being collected, instead of general policy language.

Stage 2: Abstractive Summarization

The filtered text is summarized using a Hugging Face transformer model. Initially, **sshleifer/distilbart-cnn-12-6** was used because it is fast and lightweight. However, it did not give quality summaries and changed the words to a great extent so for higher-quality summaries, facebook/bart-large-cnn was later tested. The final code uses **facebook/bart-large-cnn** with prompt-based guidance to improve relevance.

3. Model and Library Choices

- Framework: Hugging Face Transformers
- Model Used: facebook/bart-large-cnn
- Environment: Google Colab

- Tokenizer & Model Loading: Used AutoTokenizer and AutoModelForSeq2SeqLM directly instead of the pipeline() API due to compatibility issues with the newer python version.

4. Handling Large Documents

- Large PDFs (eg: 126-page Meta Privacy Policy) were converted to TXT format before processing.
- This avoids PDF encoding issues and ensures reliable text extraction.
- The system can process very large text files by chunking input internally.

5. Privacy Risk Classification

After summarization, the output is analyzed using a weighted keyword-based scoring system:

- High Risk: financial data, biometric data, precise location, data selling
- Moderate Risk: emails, phone numbers, device identifiers, ads, third parties
- Low Risk: analytics and usage data

Based on the score, the document is labeled as: Low Privacy Risk, Moderate Privacy Risk, High Privacy Risk. This makes the output more informative and actionable for users.

6. Experimental Evaluation

The system was tested usingTata Power Privacy Policy (8 pages) and Meta (Facebook/Instagram) Privacy Policy (126 pages).

Results showed that:

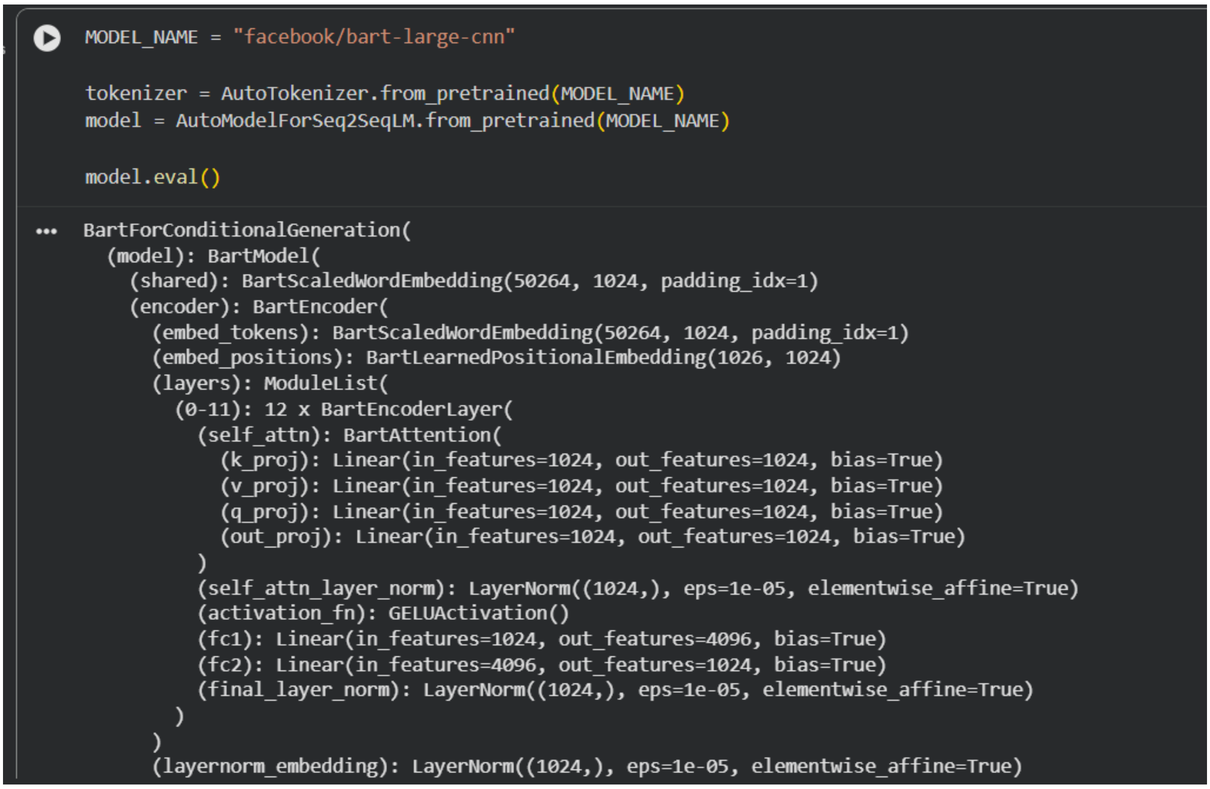
- Keyword-based extraction significantly improved relevance
- Summaries clearly highlighted types of personal data collected due to the keyword extraction used, differing from previous results which were not giving the necessary results and rather focused on the unnecessary details about the company policy.

7. Observations

- Summarization alone is insufficient for privacy analysis
- Intelligent preprocessing dramatically improves results
- Prompt engineering and keyword extraction are crucial

Output Screenshots

Model Used



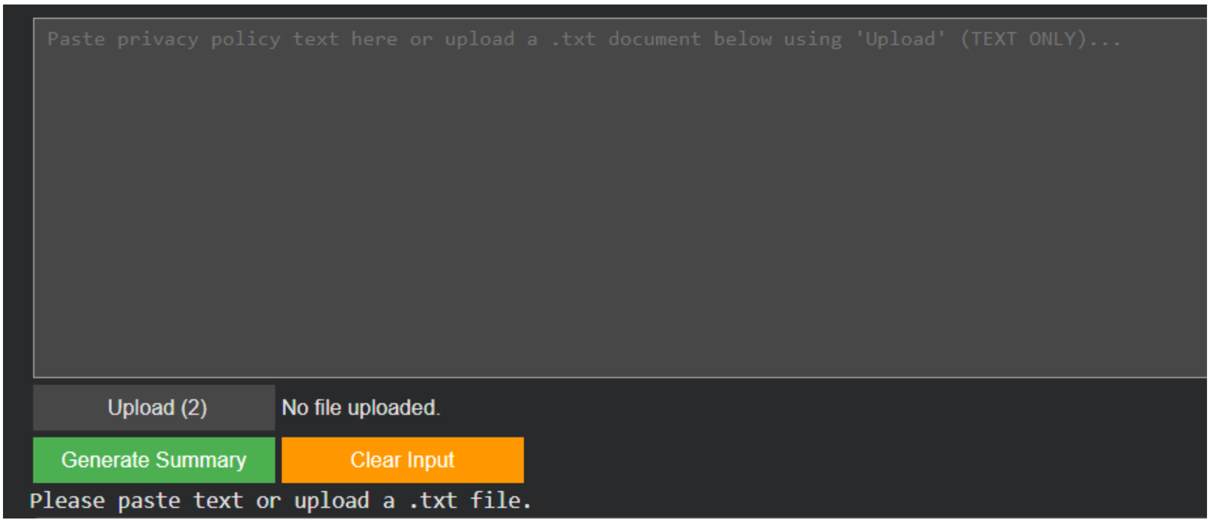
```
MODEL_NAME = "facebook/bart-large-cnn"

tokenizer = AutoTokenizer.from_pretrained(MODEL_NAME)
model = AutoModelForSeq2SeqLM.from_pretrained(MODEL_NAME)

model.eval()
```

```
BartForConditionalGeneration(
  (model): BartModel(
    (shared): BartScaledWordEmbedding(50264, 1024, padding_idx=1)
    (encoder): BartEncoder(
      (embed_tokens): BartScaledWordEmbedding(50264, 1024, padding_idx=1)
      (embed_positions): BartLearnedPositionalEmbedding(1026, 1024)
      (layers): ModuleList(
        (0-11): 12 x BartEncoderLayer(
          (self_attn): BartAttention(
            (k_proj): Linear(in_features=1024, out_features=1024, bias=True)
            (v_proj): Linear(in_features=1024, out_features=1024, bias=True)
            (q_proj): Linear(in_features=1024, out_features=1024, bias=True)
            (out_proj): Linear(in_features=1024, out_features=1024, bias=True)
          )
          (self_attn_layer_norm): LayerNorm((1024,), eps=1e-05, elementwise_affine=True)
          (activation_fn): GELUActivation()
          (fc1): Linear(in_features=1024, out_features=4096, bias=True)
          (fc2): Linear(in_features=4096, out_features=1024, bias=True)
          (final_layer_norm): LayerNorm((1024,), eps=1e-05, elementwise_affine=True)
        )
      )
      (layernorm_embedding): LayerNorm((1024,), eps=1e-05, elementwise_affine=True)
```

Before uploading or pasting any policy

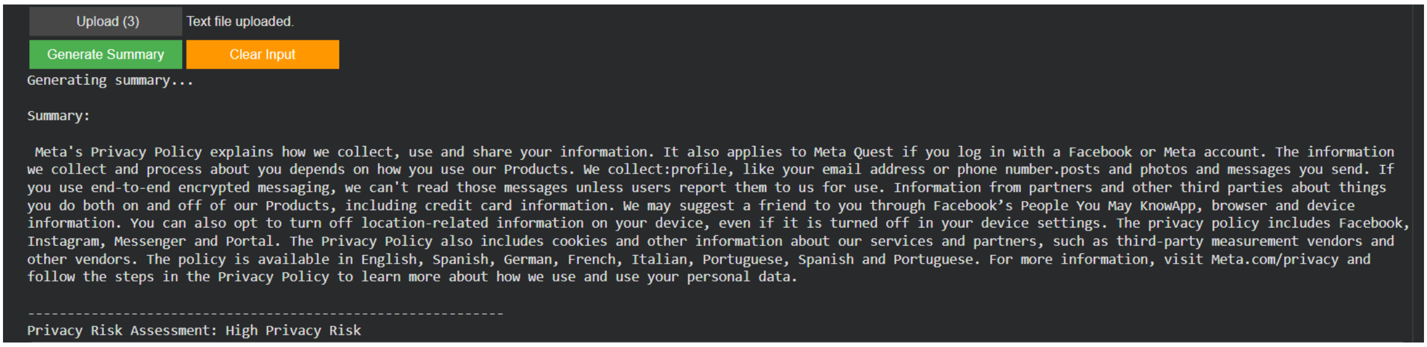


Paste privacy policy text here or upload a .txt document below using 'Upload' (TEXT ONLY)...

No file uploaded.

Please paste text or upload a .txt file.

Meta Privacy Policy Handling (126 PDF pages converted to .txt and uploaded)



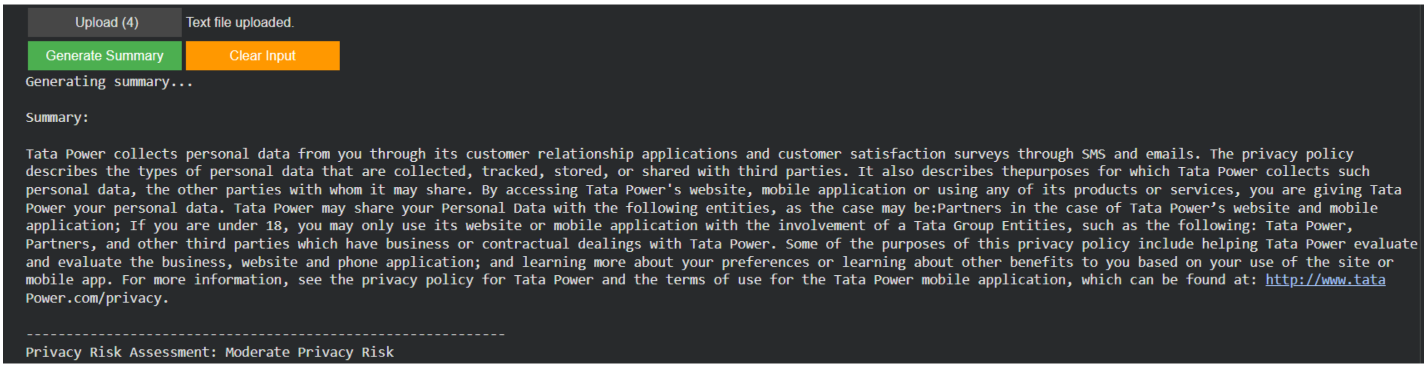
```
Generating summary...

Summary:

 Meta's Privacy Policy explains how we collect, use and share your information. It also applies to Meta Quest if you log in with a Facebook or Meta account. The information
we collect and process about you depends on how you use our Products. We collect:profile, like your email address or phone number.posts and photos and messages you send. If
you use end-to-end encrypted messaging, we can't read those messages unless users report them to us for use. Information from partners and other third parties about things
you do both on and off of our Products, including credit card information. We may suggest a friend to you through Facebook's People You May KnowApp, browser and device
information. You can also opt to turn off location-related information on your device, even if it is turned off in your device settings. The privacy policy includes Facebook,
Instagram, Messenger and Portal. The Privacy Policy also includes cookies and other information about our services and partners, such as third-party measurement vendors and
other vendors. The policy is available in English, Spanish, German, French, Italian, Portuguese, Spanish and Portuguese. For more information, visit Meta.com/privacy and
follow the steps in the Privacy Policy to learn more about how we use and use your personal data.

------------------------------------------------------------
Privacy Risk Assessment: High Privacy Risk
```

Tata Power privacy policy (8 PDF pages converted to .txt and uploaded)



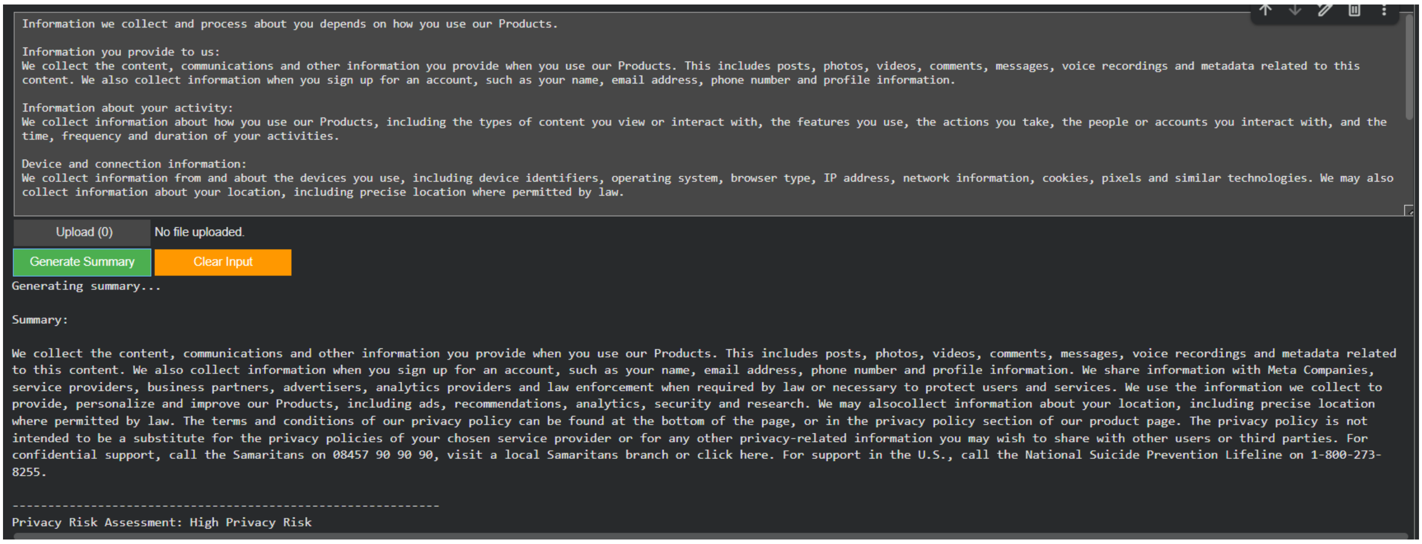
```
Generating summary...

Summary:

Tata Power collects personal data from you through its customer relationship applications and customer satisfaction surveys through SMS and emails. The privacy policy
describes the types of personal data that are collected, tracked, stored, or shared with third parties. It also describes thepurposes for which Tata Power collects such
personal data, the other parties with whom it may share. By accessing Tata Power's website, mobile application or using any of its products or services, you are giving Tata
Power your personal data. Tata Power may share your Personal Data with the following entities, as the case may be:Partners in the case of Tata Power's website and mobile
application; If you are under 18, you may only use its website or mobile application with the involvement of a Tata Group Entities, such as the following: Tata Power,
Partners, and other third parties which have business or contractual dealings with Tata Power. Some of the purposes of this privacy policy include helping Tata Power evaluate
and evaluate the business, website and phone application; and learning more about your preferences or learning about other benefits to you based on your use of the site or
mobile app. For more information, see the privacy policy for Tata Power and the terms of use for the Tata Power mobile application, which can be found at: http://www.tata
Power.com/privacy.

------------------------------------------------------------
Privacy Risk Assessment: Moderate Privacy Risk
```

Pasting Privacy Policy in the input text box



```
Information we collect and process about you depends on how you use our Products.

Information you provide to us:
We collect the content, communications and other information you provide when you use our Products. This includes posts, photos, videos, comments, messages, voice recordings and metadata related to this
content. We also collect information when you sign up for an account, such as your name, email address, phone number and profile information.

Information about your activity:
We collect information about how you use our Products, including the types of content you view or interact with, the features you use, the actions you take, the people or accounts you interact with, and the
time, frequency and duration of your activities.

Device and connection information:
We collect information from and about the devices you use, including device identifiers, operating system, browser type, IP address, network information, cookies, pixels and similar technologies. We may also
collect information about your location, including precise location where permitted by law.
```

```
Generating summary...

Summary:

We collect the content, communications and other information you provide when you use our Products. This includes posts, photos, videos, comments, messages, voice recordings and metadata related
to this content. We also collect information when you sign up for an account, such as your name, email address, phone number and profile information. We share information with Meta Companies,
service providers, business partners, advertisers, analytics providers and law enforcement when required by law or necessary to protect users and services. We use the information we collect to
provide, personalize and improve our Products, including ads, recommendations, analytics, security and research. We may alsocollect information about your location, including precise location
where permitted by law. The terms and conditions of our privacy policy can be found at the bottom of the page, or in the privacy policy section of our product page. The privacy policy is not
intended to be a substitute for the privacy policies of your chosen service provider or for any other privacy-related information you may wish to share with other users or third parties. For
confidential support, call the Samaritans on 08457 90 90 90, visit a local Samaritans branch or click here. For support in the U.S., call the National Suicide Prevention Lifeline on 1-800-273-
8255.

------------------------------------------------------------
Privacy Risk Assessment: High Privacy Risk
```