PriBot

A Chatbot for Privacy Policies

Hamza Harkous, Kassem Fawaz, Rémi Lebret, Florian Schaub, Kang G. Shin, Karl Aberer
Problem?
CHAPTER 1

Amy watched the rain through the window, her feet tapping to the rhythm of the drops. She thought about the storm outside and how it made her feel. The sound of thunder was almost comforting, reminding her of when she was a child and her mother would read stories by the window. Now, as an adult, she found solace in the rain, and it brought a smile to her face.

But as the rain began to subside, her thoughts drifted back to the past. She remembered the day she first met him, how her heart skipped a beat when he walked into the room. Every time she thought about him, her mind was filled with memories, some happy, some not so happy. She knew that he was the love of her life, but sometimes, she wondered if they were meant to be together.

Just then, the door opened, and he walked in. His smile was as warm as the rain outside, and it was like a beacon in her heart. She knew then that she would never forget him, no matter what the future held.
Solution?

Let’s turn them to a QA conversation?
UI-Limited Interfaces: Voice-activated Devices
What can we do with the current machinery?

- Read the whole policy?
What can we do with the current machinery?
What can we do with the current machinery?

Rely on secondary devices with screen?
UI-Limited Interfaces: Voice-activated Devices

You can now use any Alexa skill without enabling it first.

You no longer have to manually enable Alexa skills before you use them. Just tell Alexa to open a new skill and it will find, enable and open it with a single command.
Automated Customer Support
Automated QA Approach

Question

Policy
Automated QA Approach
Automated QA Approach
Automated QA Approach

Question
Policy

Segmenter
A1
A2
...

Privacy QA Ranking Algorithm
Q
A1
A2
...

QA Interface:
Chatbot, Voice Assistant, Twitter Bot

A4
A11
A28
A question about a policy is first segmented into parts. These segments are then ranked by a privacy QA ranking algorithm. The top-ranked answers (A4, A11, A28) are then used to generate the QA interface, which can be a chatbot, voice assistant, or Twitter bot.
Policy

Question

Segmenter

Privacy QA Ranking Algorithm

QA Interface: Chatbot, Voice Assistant, Twitter Bot
To whom do you expose my content?
To whom do you expose my content?

1. **User** wording is different from **policies** wording.
To whom do you expose my content?

1. **User** wording is different from **policies** wording.

2. Difficulty of accounting for the **general topic**:
   - Is "content" about the third parties or the first party?
Advantage of Word Embeddings
Advantage of Word Embeddings

Using a **general** embeddings, such as **GloVe** embeddings (Wikipedia14 + Gigaword 5), allows matching **words** in the **policies** to **words** used by **users**.
Neural Networks feed on labelled data.
Neural Networks feed on labelled data.

How to get such data?
Neural Networks feed on labelled data..

How to get such data?

We don't have QA data.
Neural Networks feed on labelled data..

How to get such data?

We don't have QA data.

Can we survive with classification data?
You can modify information you have given us. To correct or delete information or update account settings, log into your account and follow the instructions. We make changes as soon as we can. This information may stay in our backup files. If we cannot make the changes you want, we will let you know and explain why. If you contact us requesting access to your information, we will respond within 30 days.

You can control cookies and tracking tools. To learn how to manage how we - and our vendors - use cookies and other tracking tools, please click here.

*Wilson et al., ACL 2016; usableprivacy.org/data
Online Privacy Policies Dataset

- 115 annotated policies
- 23K annotations
To whom do you expose my content?

1. **User** wording is different from **policies** wording.

2. Difficulty of accounting for the **general topic**:
   - Is "content" about the third parties or the first party?
To whom do you expose my content?

1. **User** wording is different from policies wording.

2. Difficulty of accounting for the general topic:
   - Is "content" about the third parties or the first party?
Twitter Evaluation Dataset

- Search for unbiased keywords in replies:
  - e.g., "check our privacy policy"
- Backtrack company replies to questions
Evaluation

- Predictive Accuracy
- User-perceived Utility
Predictive Accuracy
Predictive Accuracy
Predictive Accuracy

Segmenter

$A_1$

$A_2$

$\ldots$

Privacy QA Ranking Algorithm

top-3

$A_4$

$A_{11}$

$A_{28}$
Predictive Accuracy

Segmenter

A_1

A_2

...

A_1

A_2

...

Privacy QA Ranking Algorithm

Two Experts

A_5, A_{11}

A_4

A_{11}

A_{28}

top-3
Predictive Accuracy

Privacy QA Ranking Algorithm

Two Experts

Segmenter

A₁

A₂

…

A₅, A₁₁

A₄

A₁₁

A₂₈

How many questions have an expert answer in top-k?
User-Perceived Utility

• Methodology
  • Between subject study with 4 groups
User-Perceived Utility

• Methodology
  • Between subject study with 4 groups
  • 1186 participants from MTurk (15 QA pairs per user)
UX: A Key to Chatbots' Success
UX: A Key to Chatbots' Success

• User experience is key:
  • animations, time to answer, readability, failsafe,
UX: A Key to Chatbots' Success

• User experience is key:
  • animations, time to answer, readability, failsafe,

• Balance between accuracy and usability
UX: A Key to Chatbots' Success

• User experience is key:
  • animations, time to answer, readability, failsafe,

• Balance between accuracy and usability
UX: A Key to Chatbots' Success

• User experience is key:
  • animations, time to answer, readability, failsafe,

• Balance between accuracy and usability

• Not everything has to be DL-based:
  • DL for the core functionality
  • External framework for managing interactions
Take-aways
Take-aways

• Limited-UI devices and hands-free interactions
  • Traditional privacy notice delivery methods do not apply
Take-aways

• Limited-UI devices and hands-free interactions
  • Traditional privacy notice delivery methods do not apply

• Solution: PriBot
  • Answers, automatically, user free-form question from policies
  • Provides answers that have high accuracy and relevance in real-time
Take-aways

• Limited-UI devices and hands-free interactions
  • Traditional privacy notice delivery methods do not apply

• **Solution: PriBot**
  • Answers, **automatically**, user free-form question from policies
  • Provides answers that have high accuracy and relevance in real-time

• Applications:
  • Compare privacy practices of different companies
  • Use for privacy-related customer service
Questions/Feedback?

hamzaharkous.com
hamza.harkous@gmail.com