**PriBot:** A Chatbot for Privacy Policies

Privacy policies are long and boring.

PriBot turns them into a conversation enabled with Deep Learning.

- **Limited approaches at fixing usability**
- **14.1 school years needed for understanding**
- **201 hours to read them per year**

PriBot in Action

**Do you gather my mobility?**

Your device may have sensors that provide information to assist in a better understanding of your location.

**Will you throw me under the NSA bus?**

Our legal team reviews each and every request, regardless of type, and we frequently push back when the requests appear to be overly broad or don't follow the correct process.

**Do you give my data to other companies?**

We may share non-personally identifiable information publicly and with our partners like publishers, advertisers, or connected sites. For example, we may share information publicly to show trends about the general use of our services.

PriBot QA Model: A Hierarchy of Neural Classification Networks

- **Question + Policy:**
  - **Segmenter**
  - **PriBot QA Ranking Algorithm**
  - **QA Interface:** Chatbot, Voice Assistant, Twitter Bot

Category Labels

- **1st Party Collection**
- **3rd Party Sharing**

Value Labels

- **Financial**
- **Location**
- **Marketing**
- **Legal**
- **On-website**
- **From 3rd Party**

Distribution vector for: 

\[ P(\text{category}) \cdot P(\text{value}) \]

**Similarity Computation**

**Answer Vector**

**Fr주일**

**Text Segment**

- **Frozen Embeddings Layer (GloVe)**
- **CNN + Max-pooling**
- **Dense Layer + Relu**
- **Dense Layer**

User Perceived Utility

- 91% of participants found a satisfactory answer in top-3
- 55% of users found the top-answer satisfactory
- 10 users evaluated each QA pair
- 1,186 participants

Expert-based Predictive Accuracy

- 64% of questions had an answer matching experts’ answers in top-3
- 2 experts produced answers
- 10% accuracy boost over retrieval or traditional neural QA
- Higher improvement margins for longer policies

**Law Experts’ Annotations**

- OPP-15 dataset from Usable Privacy Project

**Usable policies on emerging devices**

- Limited approaches at fixing usability
- 14.1 school years needed for understanding
- 201 hours to read them per year

**Simplified policies in a chat medium**

- Automated customer privacy service
- Pretrained no need for per-company training data
- Automated customer privacy service

201 hours to read them per year

14.1 school years needed for understanding

Limited approaches at fixing usability

**PriBot in Action**

- **Do you gather my mobility?**
  - Your device may have sensors that provide information to assist in a better understanding of your location.
- **Will you throw me under the NSA bus?**
  - Our legal team reviews each and every request, regardless of type, and we frequently push back when the requests appear to be overly broad or don't follow the correct process.
- **Do you give my data to other companies?**
  - We may share non-personally identifiable information publicly and with our partners like publishers, advertisers, or connected sites. For example, we may share information publicly to show trends about the general use of our services.

**PriBot QA Model: A Hierarchy of Neural Classification Networks**

- **Question + Policy:**
  - **Segmenter**
  - **PriBot QA Ranking Algorithm**
  - **QA Interface:** Chatbot, Voice Assistant, Twitter Bot

**Category Labels**

- **1st Party Collection**
- **3rd Party Sharing**

**Value Labels**

- **Financial**
- **Location**
- **Marketing**
- **Legal**
- **On-website**
- **From 3rd Party**

**Distribution vector for:**

\[ P(\text{category}) \cdot P(\text{value}) \]

**Similarity Computation**

**Answer Vector**

**Fr주일**

**Text Segment**

- **Frozen Embeddings Layer (GloVe)**
- **CNN + Max-pooling**
- **Dense Layer + Relu**
- **Dense Layer**

**User Perceived Utility**

- 91% of participants found a satisfactory answer in top-3
- 55% of users found the top-answer satisfactory
- 10 users evaluated each QA pair
- 1,186 participants

**Expert-based Predictive Accuracy**

- 64% of questions had an answer matching experts’ answers in top-3
- 2 experts produced answers
- 10% accuracy boost over retrieval or traditional neural QA
- Higher improvement margins for longer policies