PriBot

A Chatbot for Privacy Policies

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

Hamza Harkous, Kassem Fawaz, Rémi Lebret,
Florian Schaub, Kang G. Shin, Karl Aberer

UNIVERSITY OF MICHIGAN

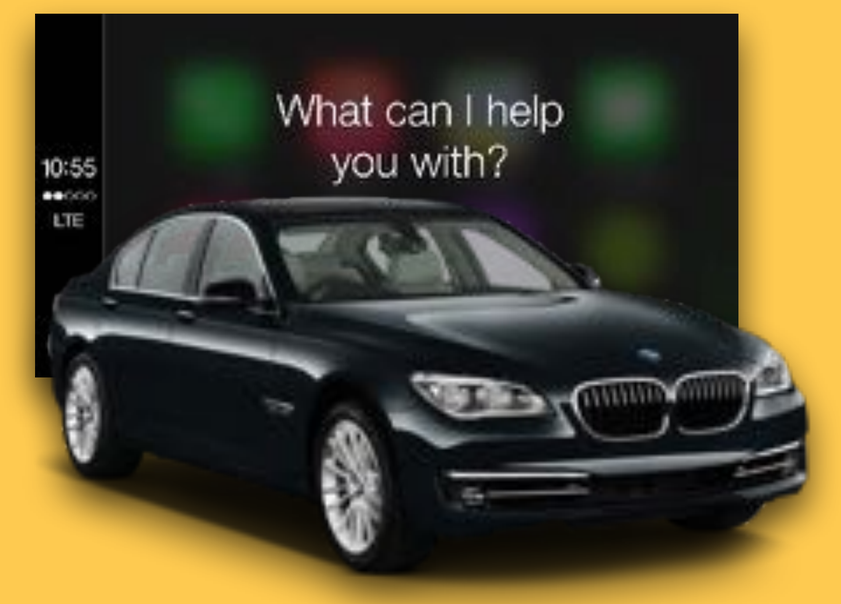
Problem?


CHAPTER 1
A late summer breeze wafted through the open kitchen window, making the twenty tiny flames upon Cecily's cake dance back and forth on their candlewicks. Cecily hadn't made the cake, of course, as one should never bake her own birthday cake, but her mother was a good cook and a better baker, so Cecily had no doubts that the confection, complete with pink cherry frosting and jelly filling, would be delicious.
But as her parents and three siblings sang her birthday wishes, Cecily's mind wandered from the dessert and the celebration at hand. Her thoughts narrowed in on an image she had seen in a ferocity box just three months ago, after reading Magician Emery Thane's fortune. A flowery hill at sunset,
kindle

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?

amazon echo

Google Home

What can we do with the current machinery?

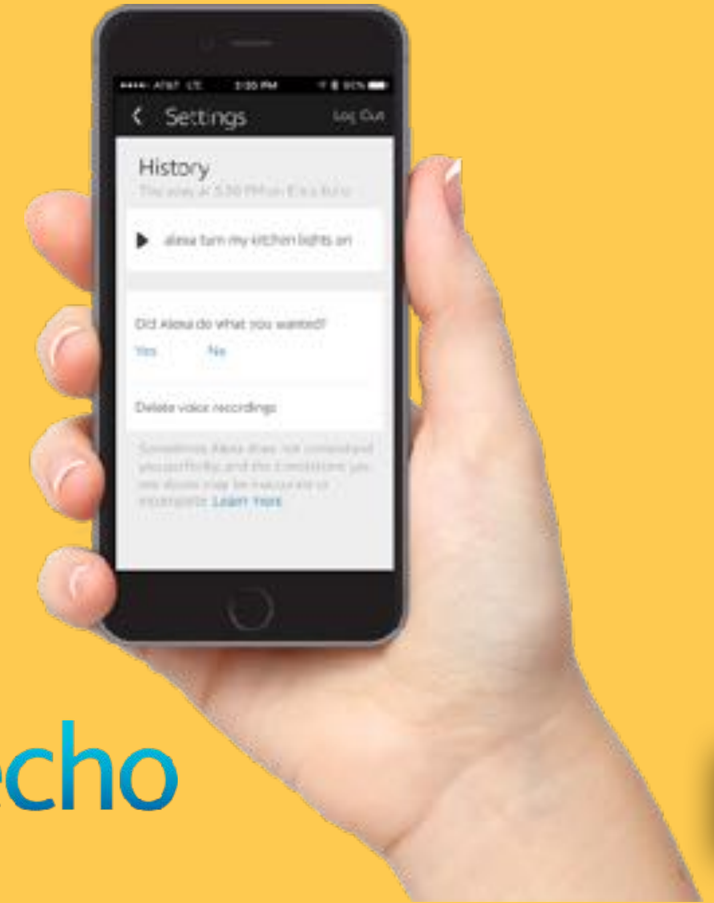
amazon echo

Google Home

What can we do with the current machinery?

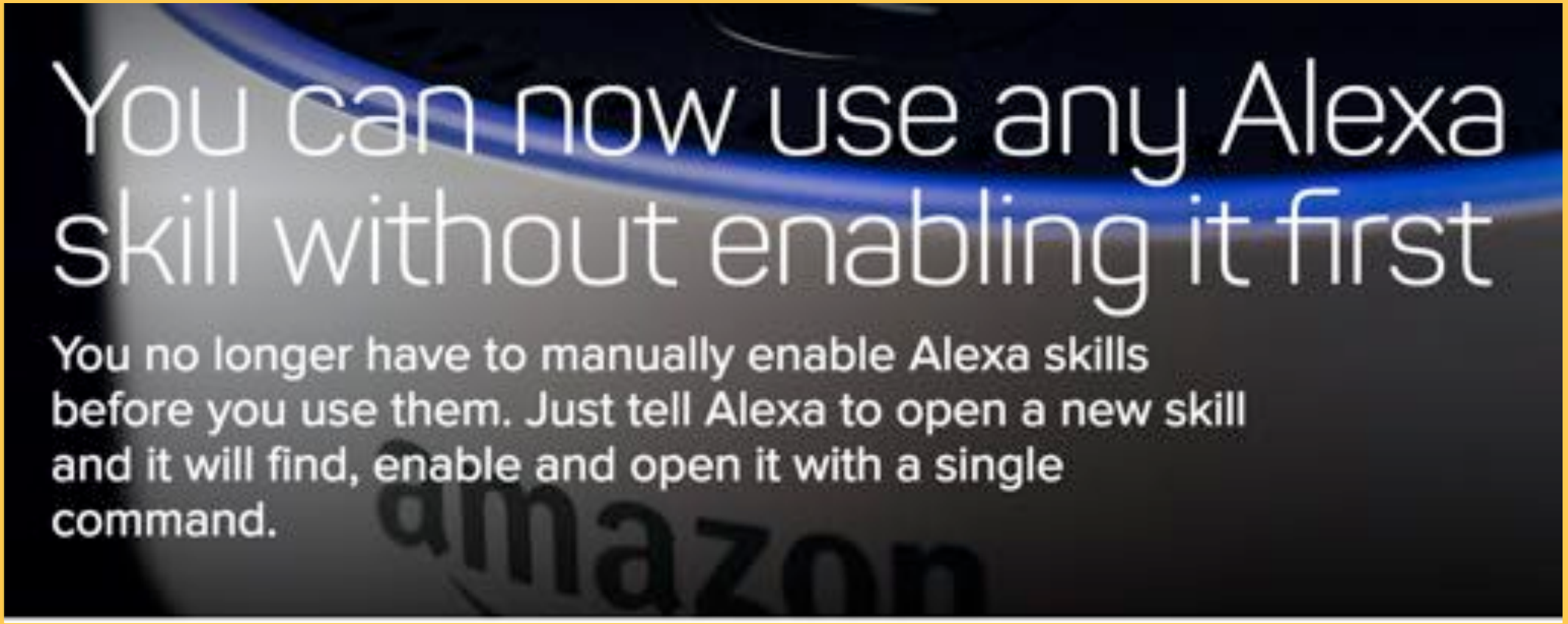
Rely on secondary devices with screen?

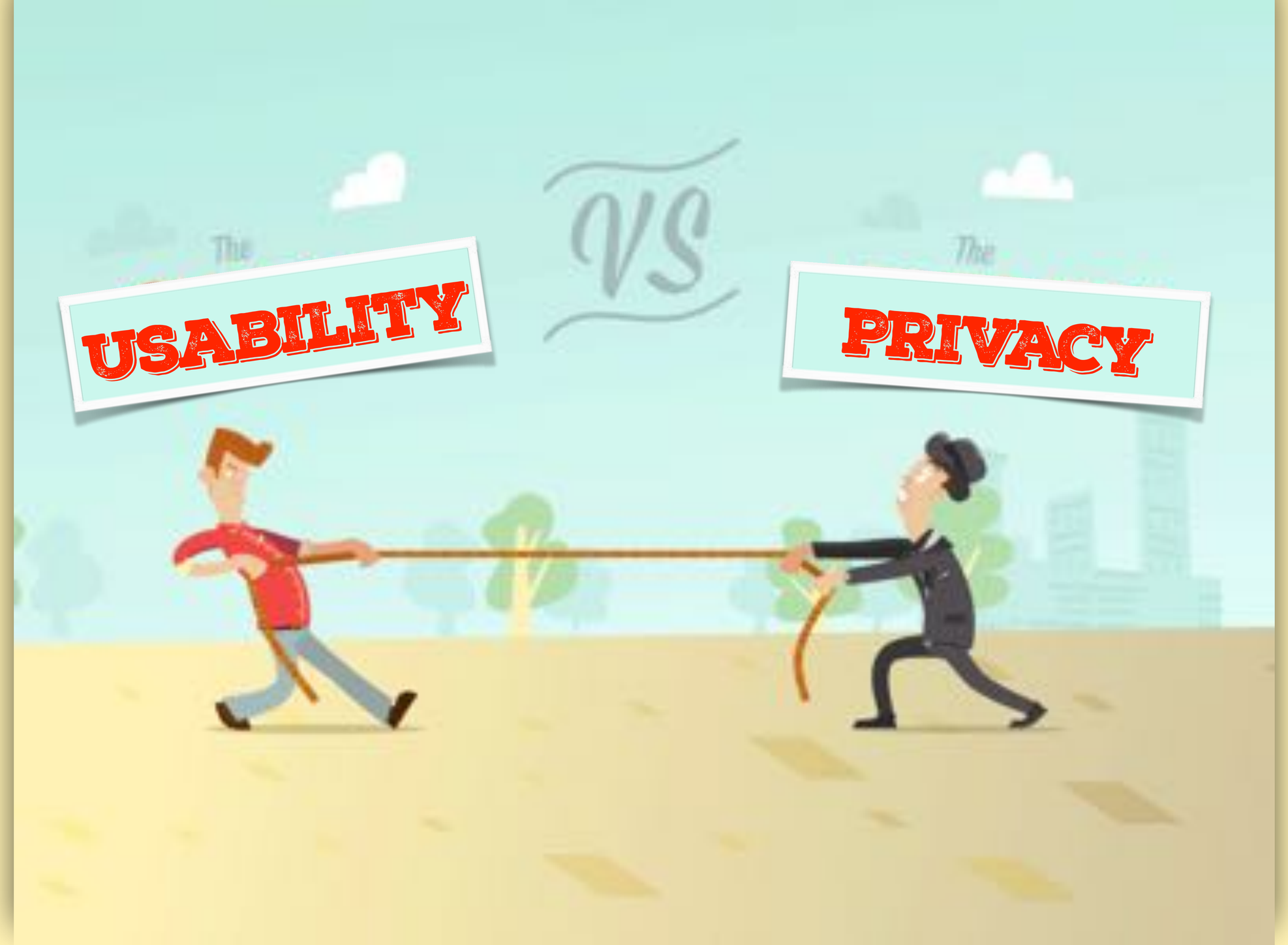
amazon echo



Google Home

UI-Limited Interfaces: Voice-activated Devices




The
USABILITY
VS
The
PRIVACY

Customer Support

Automated

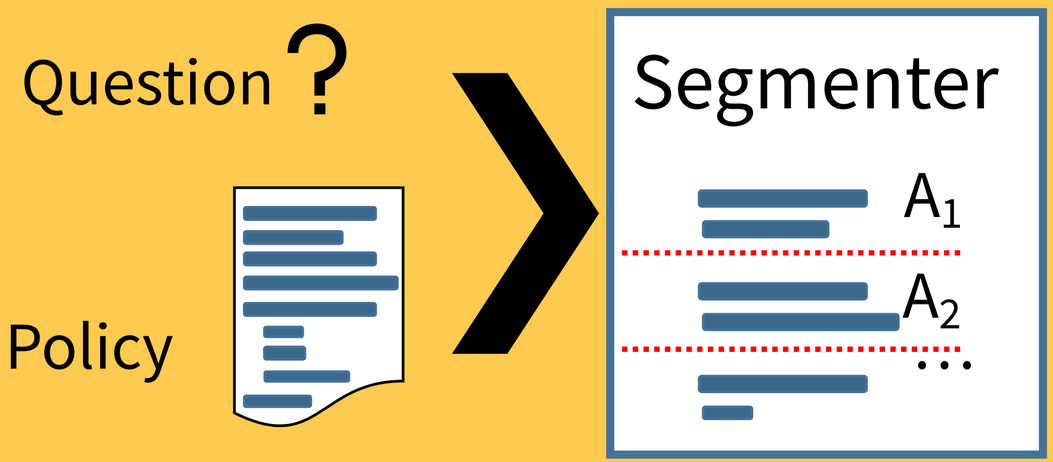
Customer Support

Automated QA Approach

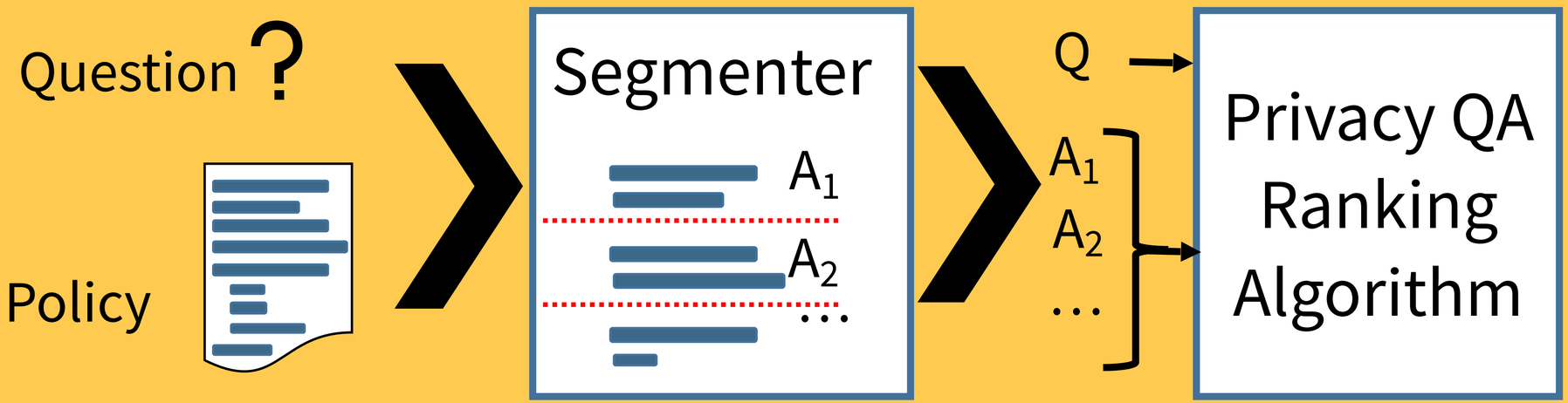
Question ?

Policy

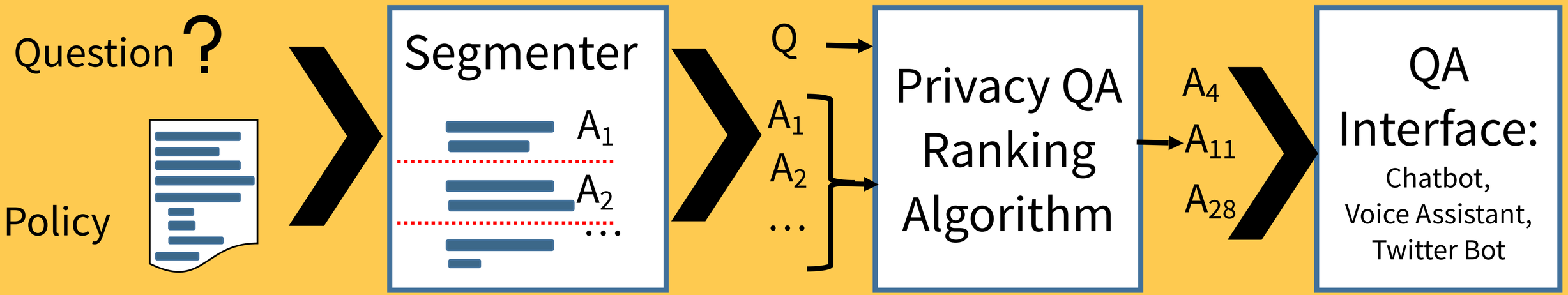
Automated QA Approach

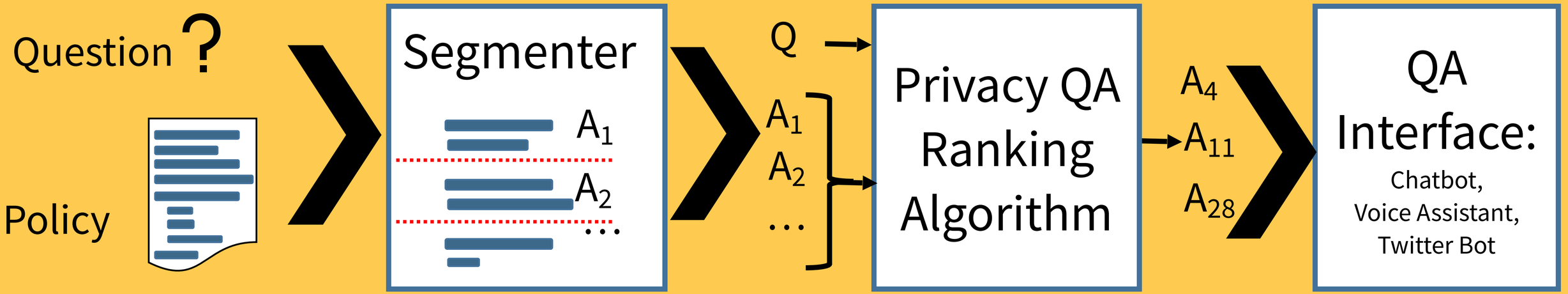


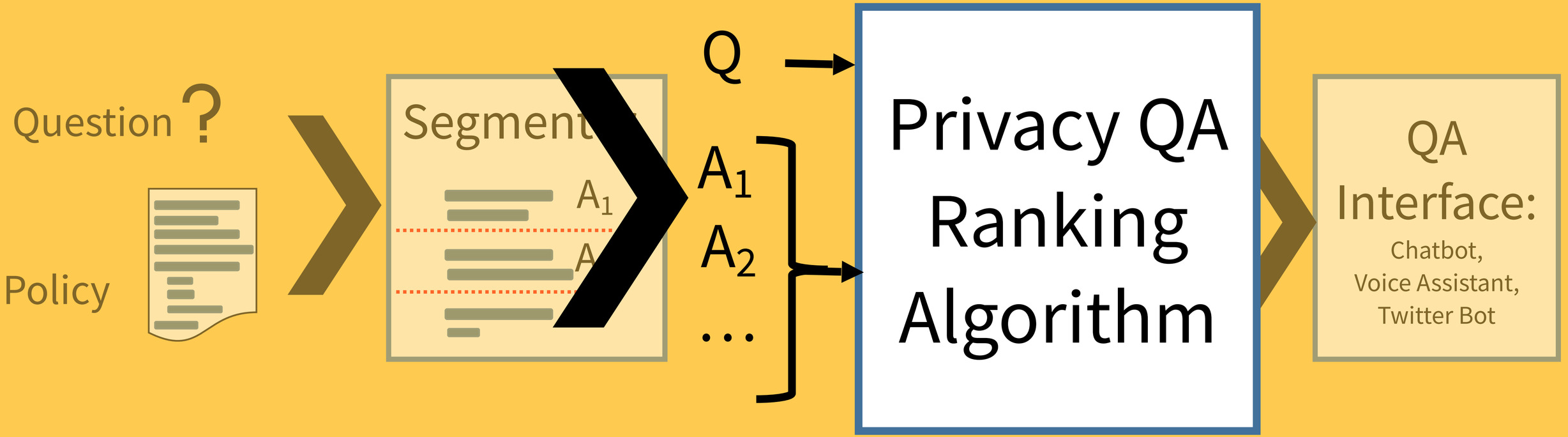
Automated QA Approach



Automated QA Approach




Question ?
Policy
Segmenter
A1
A2
...
Q
A1
A2
...
Privacy QA Ranking Algorithm
A4
A11
A28
QA Interface:
Chatbot,
Voice Assistant,
Twitter Bot


Question
Policy
Segment
A1
A
Q
A1
A2
...
Privacy QA
Ranking
Algorithm
QA
Interface:
Chatbot,
Voice Assistant,
Twitter Bot

Ranking Challenges

To whom do you expose my content?

Ranking Challenges

To whom do you expose my content?

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

Ranking Challenges

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?

Online Privacy Policies Dataset (OPP)*

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.

User Access, Edit & Deletion

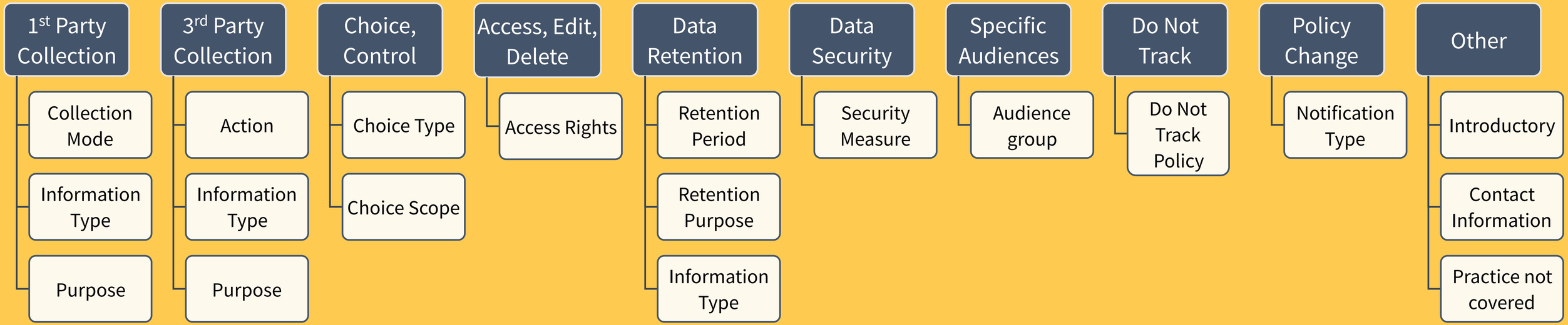
Access Type: Edit Information

Expert Annotations

*Wilson et al., ACL 2016; usableprivacy.org/data

Online Privacy Policies Dataset

- 115 annotated policies
- 23K annotations



Ranking Challenges

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?

Ranking Challenges

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

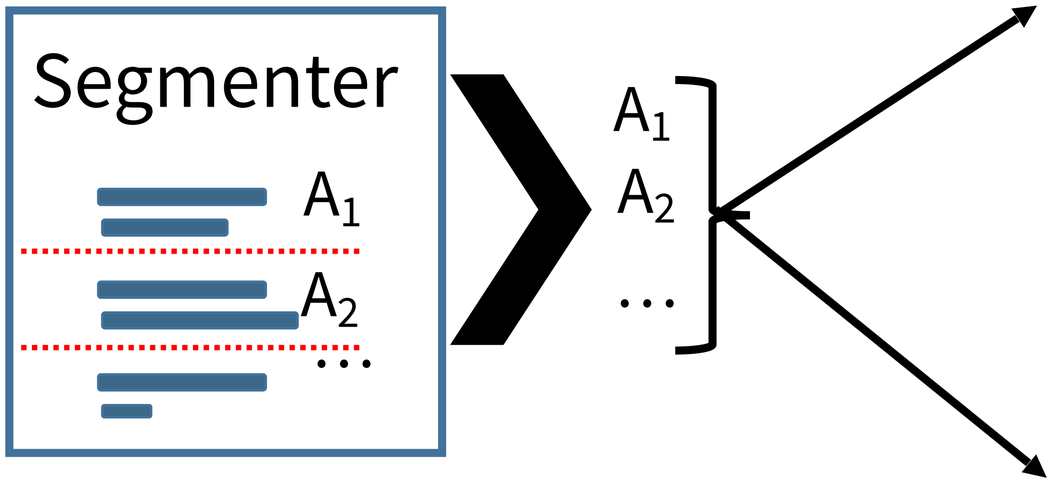
- Search for unbiased keywords in replies:
  - e.g.,: "check our privacy policy"
- Backtrack company replies to questions

Evaluation

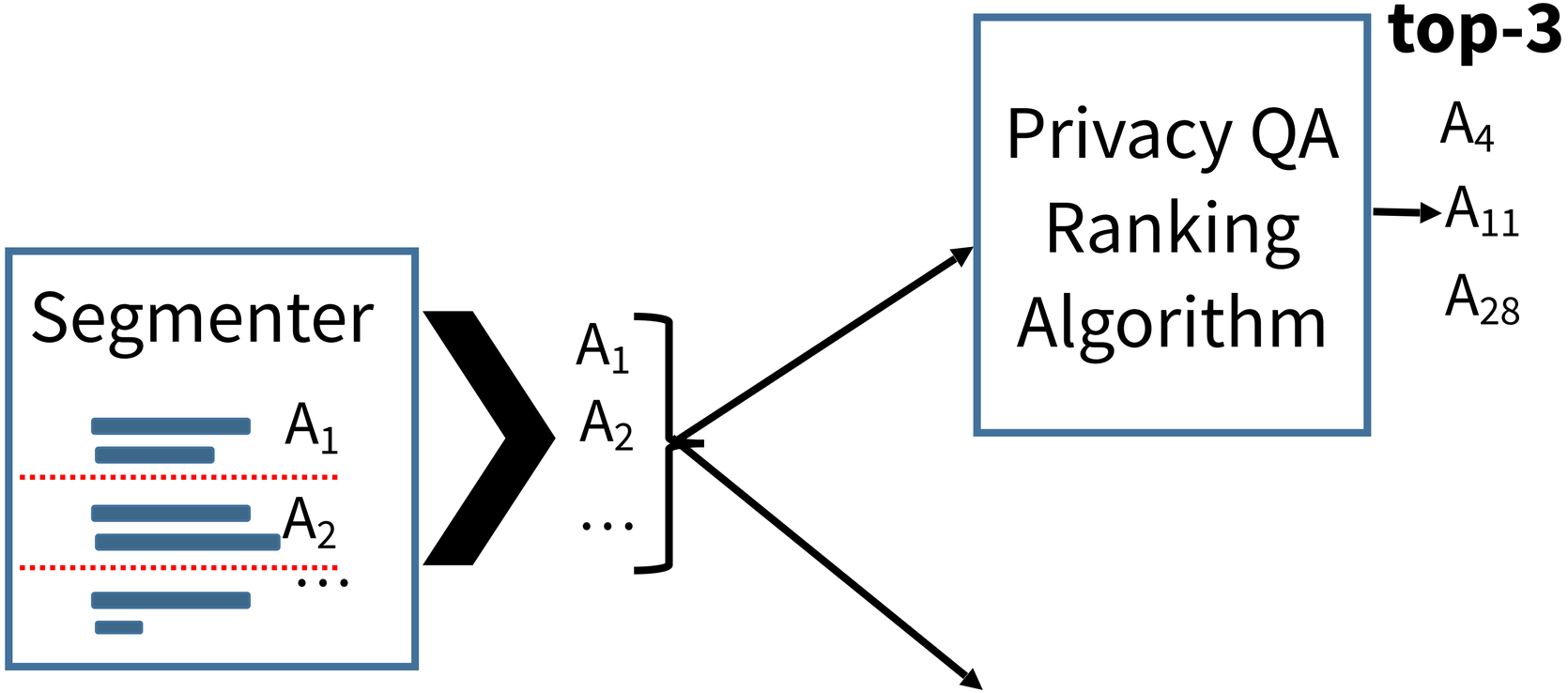
- Predictive Accuracy
- User-perceived Utility

Predictive Accuracy

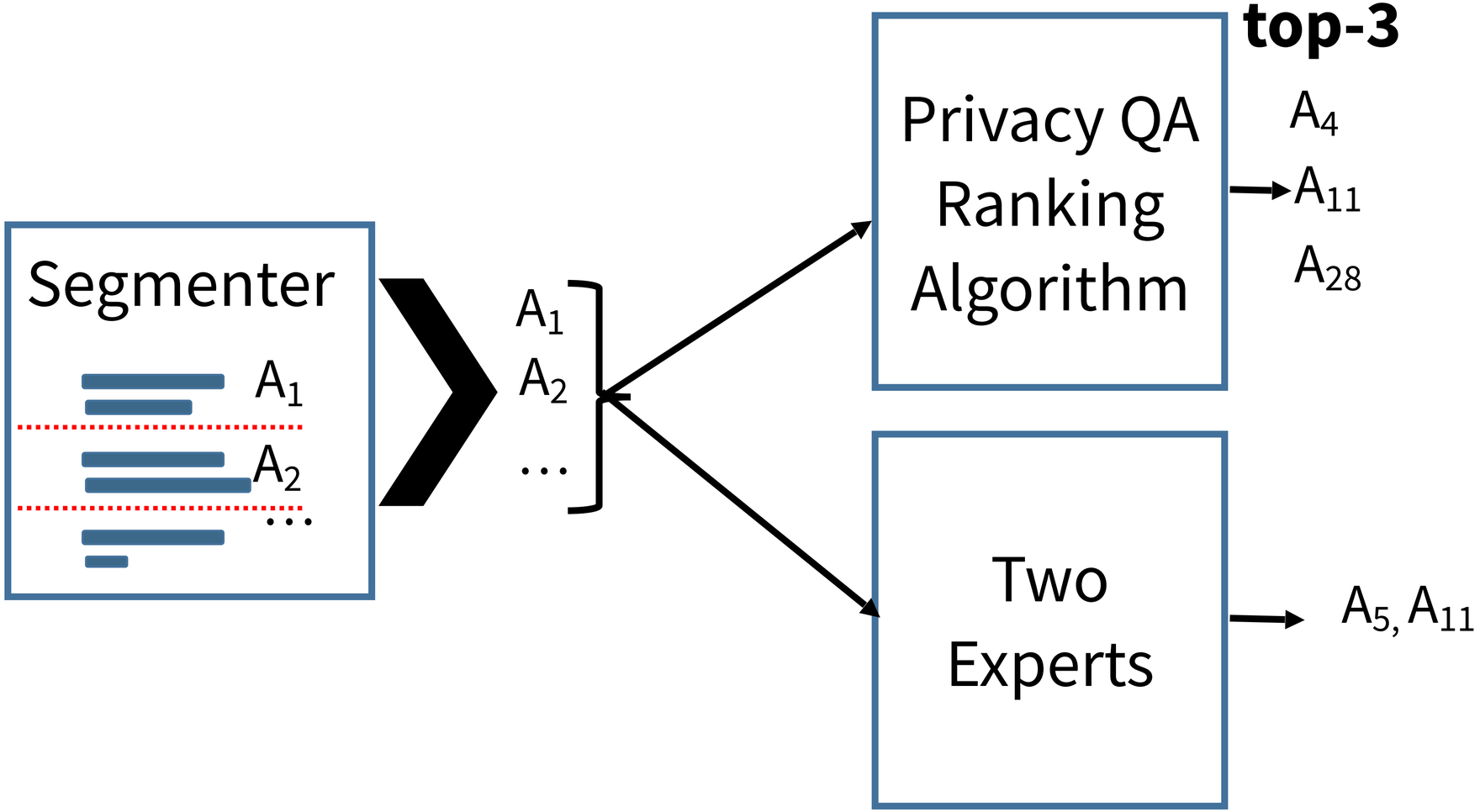
Predictive Accuracy



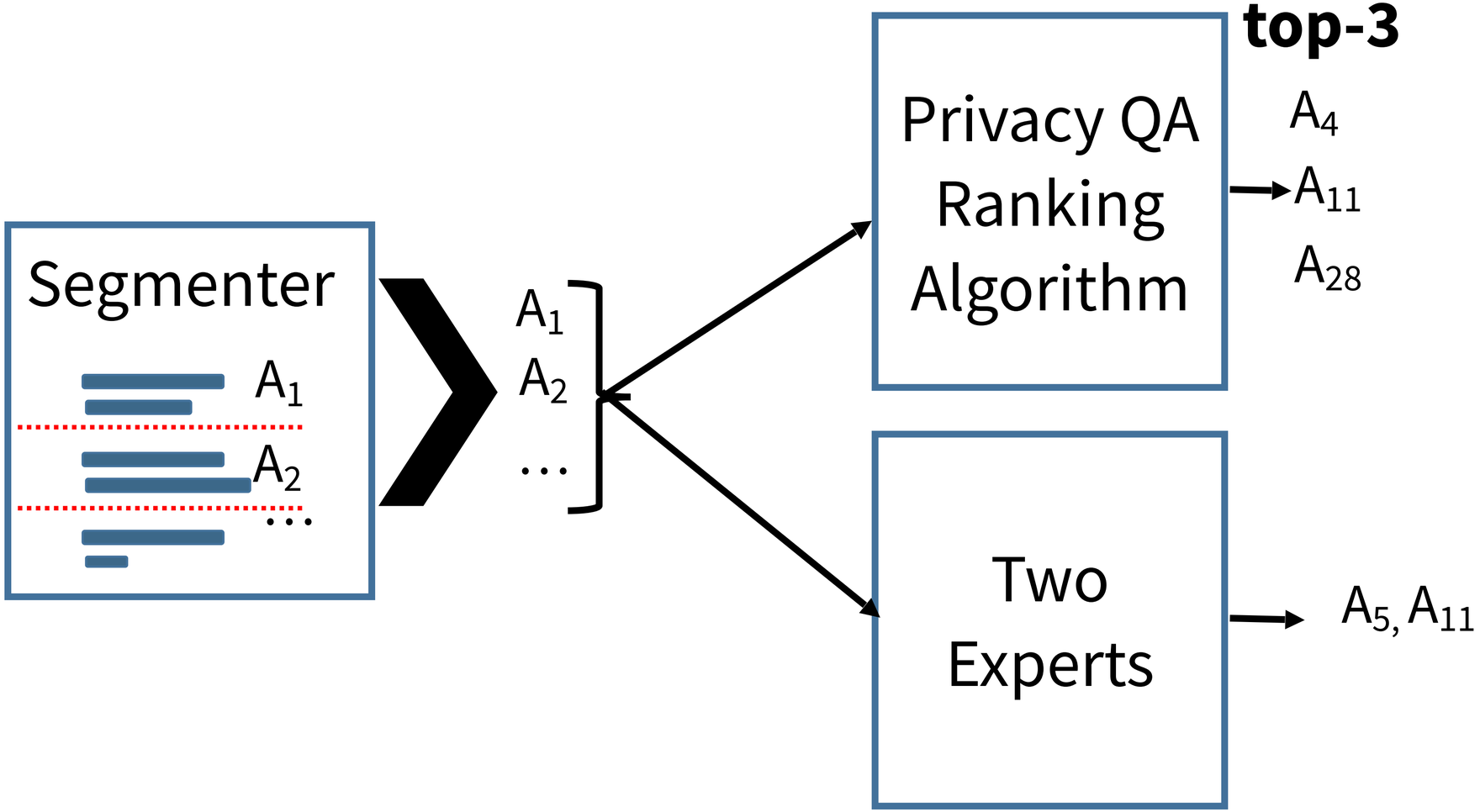
Predictive Accuracy



Predictive Accuracy



Predictive Accuracy



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