

# Detecting Textual Saliency in Privacy Policy

---



Phase-2 Progress Updates



# Goal

Provide key information to end-users using Machine Learning and Natural Language Processing



**INFORM**  
Module

- ★ Visualization
- ★ Annotating policies



**QUERY**  
Module

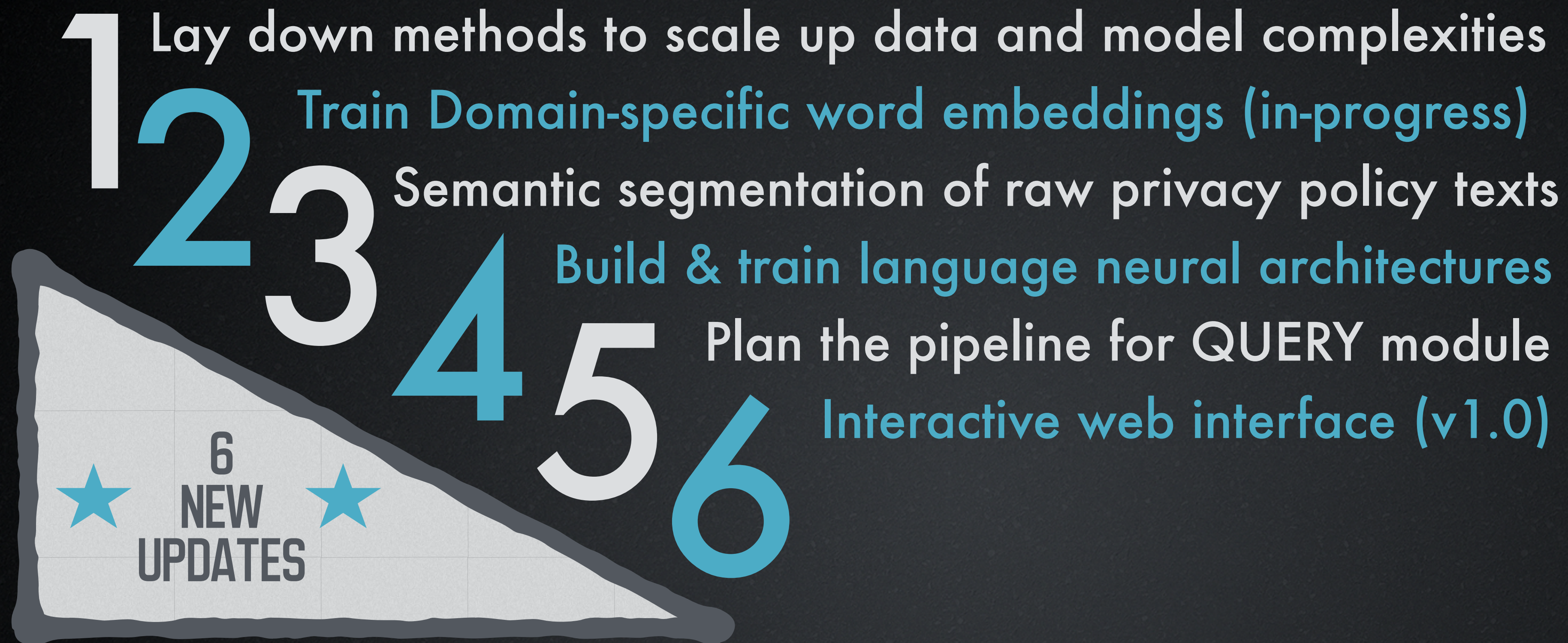
- ★ Enable users to ask policy-related

Interactive  
**Web Framework**





# Progress Updates



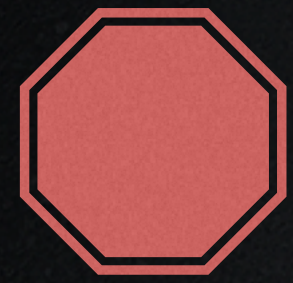
★ Major focus on the implementation of web interface



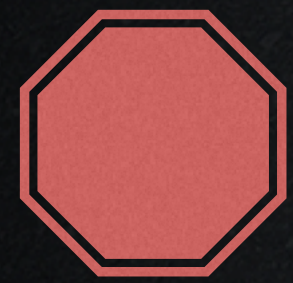
# Roadblocks



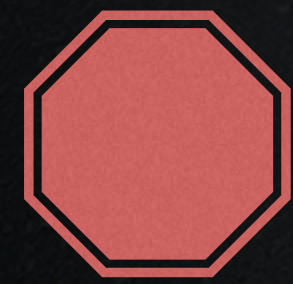
Data and model scalability issues



Semantic segmentation



Building custom BERT architectures for classification and QA



Build, deploy and integrate the model stack with the web interface



# Methodology

## (Phase-2)

- ★ Reliable predictions from baselines
- ★ Hard to predict classes with lower counts of positive examples
- ★ NEXT: Scale up data and model complexities
- ★ Understand strengths and weaknesses of each model in predicting a category
- ★ Leverage the expertise of each learner to perform predictions on unseen data



# Scaling Up Data Complexity

## (Domain-Specific Non-Contextual Embeddings)

- ★ Limited labeled data => Feed unlabeled data first, then increase model complexity
- ★ Pre-trained generic embeddings showed very little performance improvement
- ★ Previous literature recommends training domain-specific embeddings, if possible

## Dataset

- ★ Obtained access to 1 M raw privacy policies dump (PrivaSeer Corpus)
- ★ May not need all 1 M policies (performance may saturate sooner)



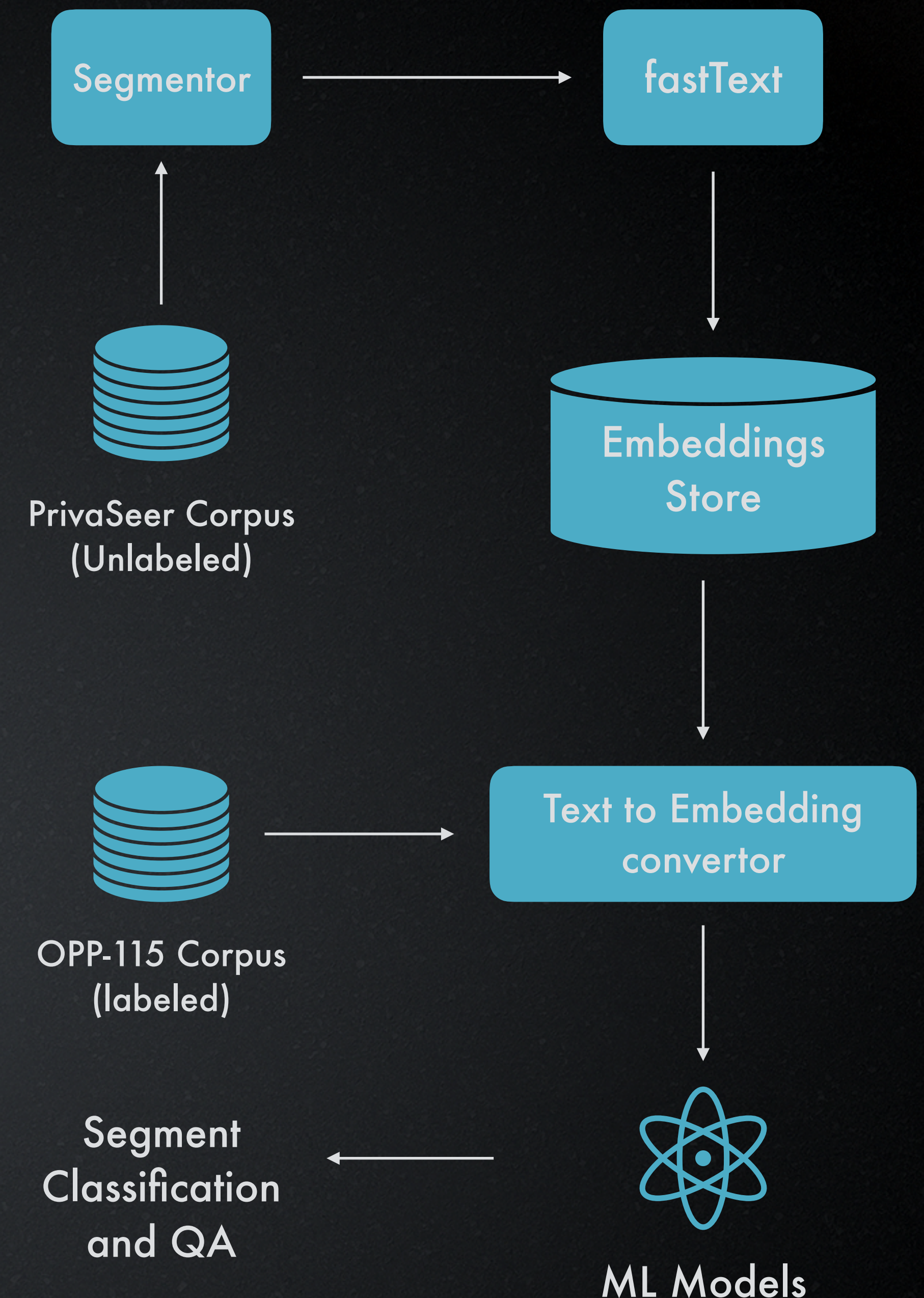
# Domain-Specific Non-Contextual Embeddings

## (Process)

- ★ Significant time spent on segmentation, fastText and handling the scale of the corpus (still in-progress)
- ★ Automatic Segment Detection using Unsupervised and Supervised Learning (ASDUS)

## fastText

- ★ Open-source library that allows users to learn supervised and unsupervised text representations and text classifiers.
- ★ Supports both continuous bag of words and the Skip-gram model
- ★ Performs better than Word2Vec and GloVe - uses character level information to handle OOV or rare words.
- ★ Our objective: train semantic embeddings for downstream tasks





# Automatic Segment Detection using Unsupervised and Supervised Learning (ASDUS)

Firefox

Internet Health

Technology

About Us

Get Involved

Download Firefox

April 15, 2014

Your privacy is an important factor that Mozilla (that's us) considers in the development of each of our products and services. We are committed to being transparent and open. This Mozilla Privacy Policy explains generally how we receive information about you, and what we do with that information once we have it. You can also read our product privacy notices and FAQs to get more detail with respect to each of our products and services.

What do we mean by "personal information?"

For us, "personal information" means information which identifies you, like your name or email address.

Any information that falls outside of this is "non-personal information."

If we store your personal information with information that is non-personal, we will consider the combination as personal information. If we remove all personal information from a set of data then the remaining is non-personal information.

How do we learn information about you?

We learn information about you when:

you give it to us directly (e.g., when you choose to send us crash reports);

we collect it automatically through our products and services (e.g., when your Firefox browser checks with us to see if is up to date);

someone else tells us information about you (e.g., when Thunderbird works with your email providers to set up your account); or

when we try and understand more about you based on information you've given to us (e.g., when we use your IP address to customize language for some of our services).

What do we do with your information once we have it?

When you give us information, we will use it in the ways for which you've given us permission. Generally, we use your information to help us provide and improve our products and services for you.

Data Privacy Principles

Mozilla is an open source project with a mission to improve your Internet experience. This is a driving force behind our data privacy practices. [Read More »](#)

Our Privacy Notices

We created short and clear Privacy Notices to describe how each of our products and services receives, shares, and uses data and what your choices are. [Learn more:](#)

[Mozilla Websites, Communications & Cookies](#)

[Firefox Browser](#)

[Firefox OS](#)

[Firefox + Cliz](#)

[Firefox Cloud Services](#)

[Firefox Marketplace](#)

[Firefox Focus](#)

[Persona](#)

## Uncategorized

April 15, 2014

Your privacy is an important factor that Mozilla (that's us) considers in the development of each of our products and services. We are committed to being transparent and open. This Mozilla Privacy Policy explains generally how we receive information about you, and what we do with that information once we have it. You can also read our product privacy notices and FAQs to get more detail with respect to each of our products and services.

### What do we mean by "personal information?"

For us, "personal information" means information which identifies you, like your name or email address.

Any information that falls outside of this is "non-personal information."

If we store your personal information with information that is non-personal, we will consider the combination as personal information. If we remove all personal information from a set of data then the remaining is non-personal information.

### How do we learn information about you?

We learn information about you when:

you give it to us directly (e.g., when you choose to send us crash reports);

we collect it automatically through our products and services (e.g., when your Firefox browser checks with us to see if is up to date);

someone else tells us information about you (e.g., when Thunderbird works with your email providers to set up your account); or

when we try and understand more about you based on information you've given to us (e.g., when we use your IP address to customize language for some of our services).

### What do we do with your information once we have it?

When you give us information, we will use it in the ways for which you've given us permission. Generally, we use your information to help us provide and improve our products and services for you.

### When do we share your information with others?

When we have asked and received your permission to share it.

For processing or providing products and services to you, but only if those entities receiving your information are contractually obligated to handle the data in ways that protect your privacy.

When we are fulfilling our . We sometimes release information to make our products better and foster an open web, but when we do so, we will remove your personal information from the data we release. We will not release your personal information in a way that could identify you, unless we have your permission or the law requires us to do so.

When the law requires it. We follow the law whenever we receive requests about you from a government or related to a lawsuit. We'll notify you when we're asked to hand over your information, unless the law requires us not to. Nothing in this policy is intended to limit any legal rights or remedies you may have.

When we believe it is necessary to prevent harm to you or someone else. We will only share your information in this way if we have a good faith belief that it is reasonable to do so to prevent or mitigate harm to you or someone else.

If our organizational structure or status changes (if we undergo a restructuring, are acquired, or go bankrupt) we may pass your information to a successor or affiliate.

```
graph TD; Start([Start]) --> HTML[/HTML file/]; HTML --> Parsing[Parsing & Text Collection]; Parsing --> Feature[Feature Extraction]; Feature --> Clustering[Clustering & Segment Identification]; Clustering --> Version[Simple Version Generation]; Version --> End([End]);
```

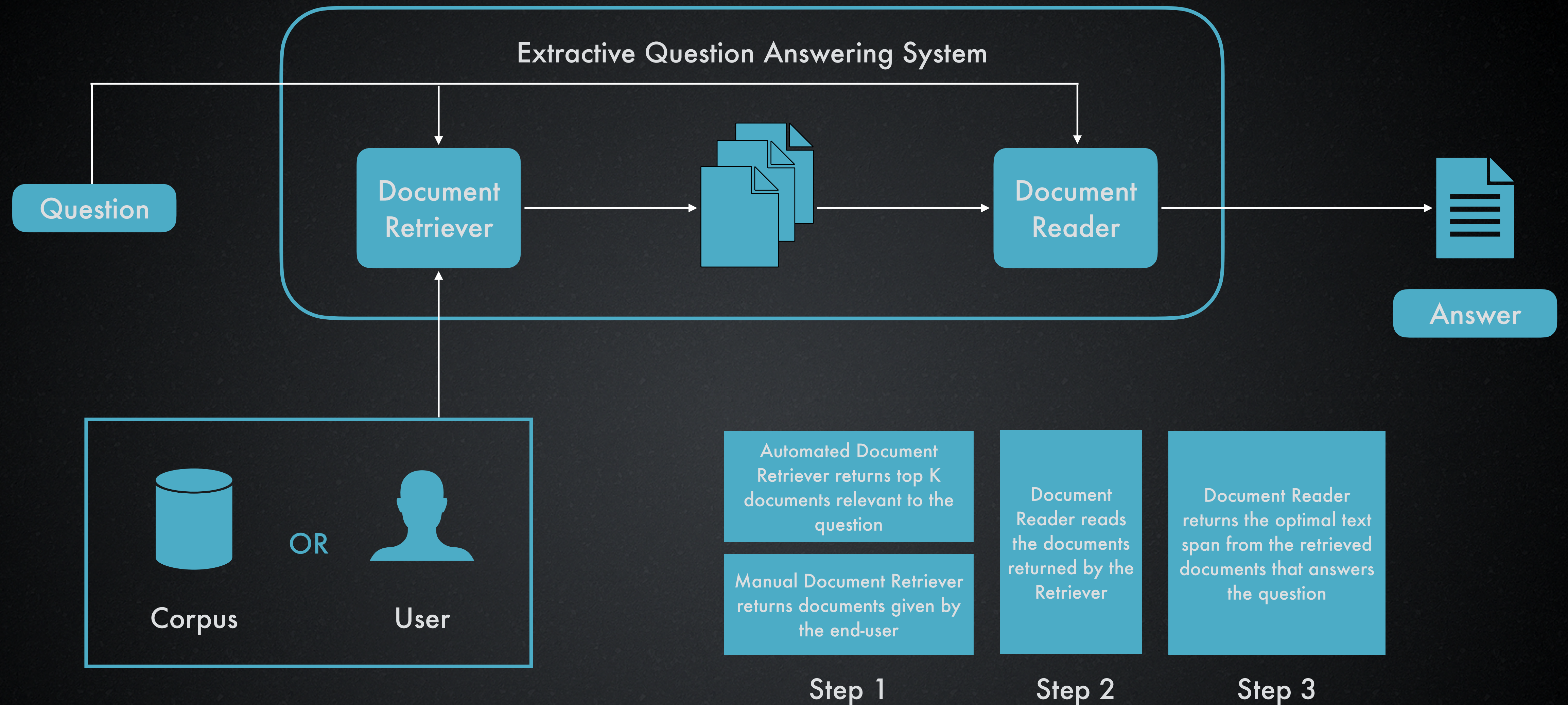
The flowchart illustrates the Automatic Segment Detection (ASDUS) process. It begins with a 'Start' node, followed by an 'HTML file' input. The process then proceeds through 'Parsing & Text Collection', 'Feature Extraction', 'Clustering & Segment Identification', and 'Simple Version Generation', finally reaching an 'End' node. Brackets on the right side of the flowchart group the steps into two categories: 'Java Code' (covering Parsing & Text Collection, Feature Extraction, and Clustering & Segment Identification) and 'Python Code' (covering Simple Version Generation).

Source



# Question-Answering System

## (QUERY Module)





# Question-Answering System

(Continued)

- ★ Extractive Question Answering
- ★ Use a recently released PolicyQA dataset curated from the OPP-115 corpus
  - ◆ Enables us to model the task as predicting the answer in the given policy segment using existing neural approaches from literature (utilizing Hugging Face transformer API)
  - ◆ A similar dataset SQuAD (Stanford Question Answering Dataset) curated from Wikipedia articles could help provide more insight into the process.
  - ◆ Use pretrained BERT on SQUAD 2.0 dataset and fine-tune it on the PolicyQA dataset (with domain-specific word embeddings)



## Example 1

Policy Text

Information You Give Us: We receive and store any **information you enter on our Web site** or give us in any other way. Click here to see ...

Question

How do you collect my information?

Answer

**information you enter on our Web site**

## Example 2

Promotional Offers: Sometimes we send offers to selected groups of Amazon.com customers on behalf of other businesses. When we do this, **we do not give that business your name and address**. If you do not want to receive such offers, ...

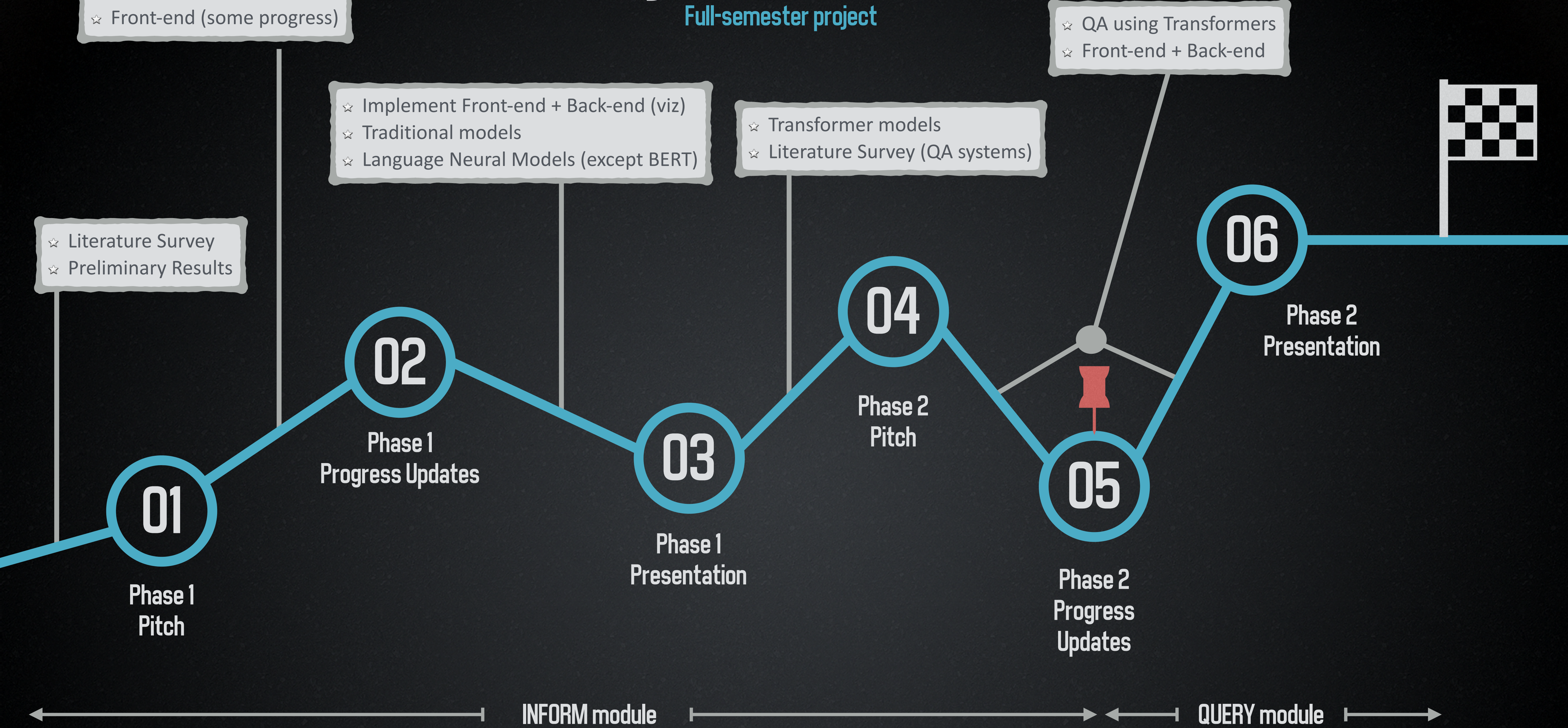
Is my information shared with others?

**we do not give that business your name and address**



# Project Timeline

Full-semester project





# Final Deliverables

(INFORM + QUERY modules)



Complete training advanced models for both the classification and QA tasks



Build and Integrate QA with the web interface



Build a model playground using Streamlit for technical audience



Document, attribute, package, host and deploy the complete project stack on Heroku



THANK  
YOU

