

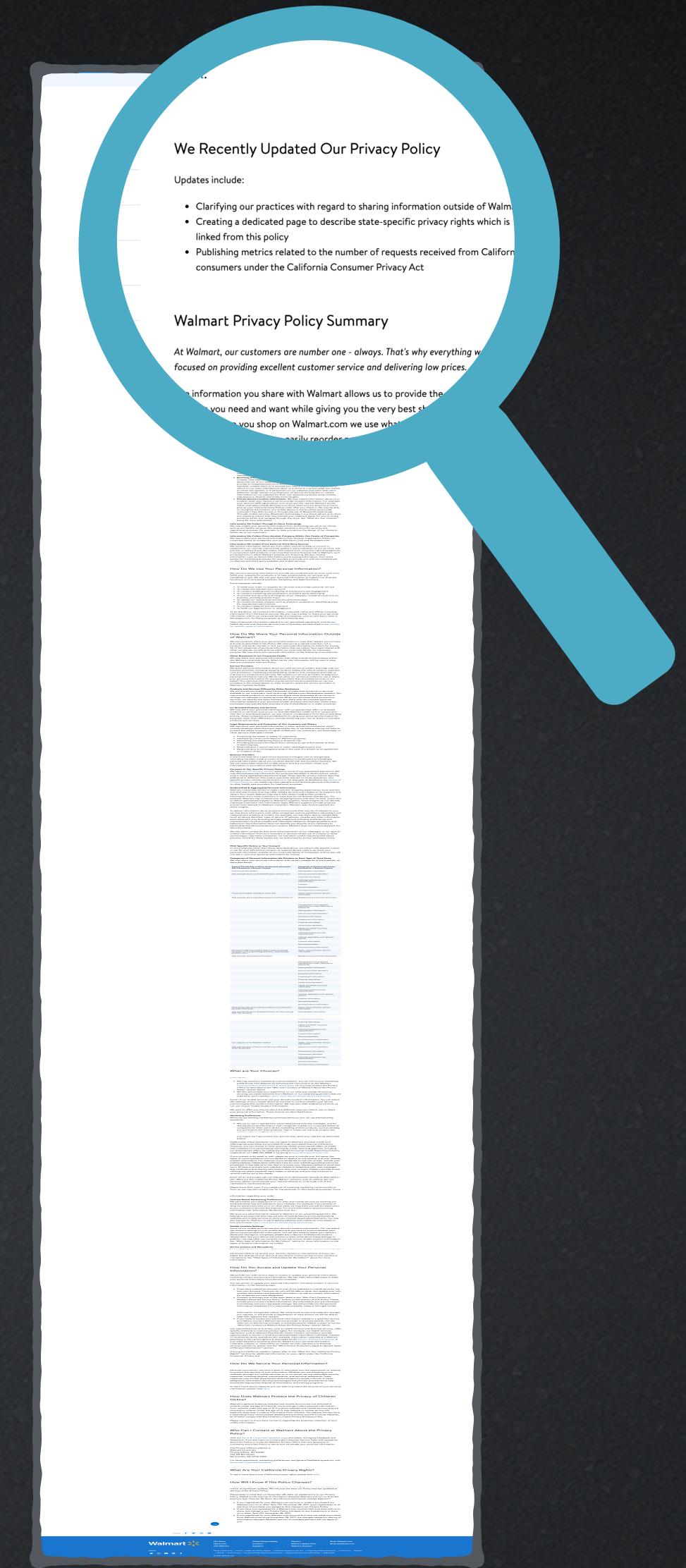
# Detecting Textual Saliency in Privacy Policy

---



Phase-1 Progress Updates

# Recap: A Close Look At Privacy Policy



## Details about

- ? How and why do they collect our information?
- ? Are they sharing this data with third parties?
- ? How long is my data stored?
- ? Can I access my data?
- ? How secure and protected is my data?

# Challenges

## 01 Increase in complexity

More than 85% of privacy policies scraped are at or above college level reading difficulty ([Source](#))

## 02 Increase in length and verbosity

Privacy policies are 1.8x longer than in 2009 ([Source](#))

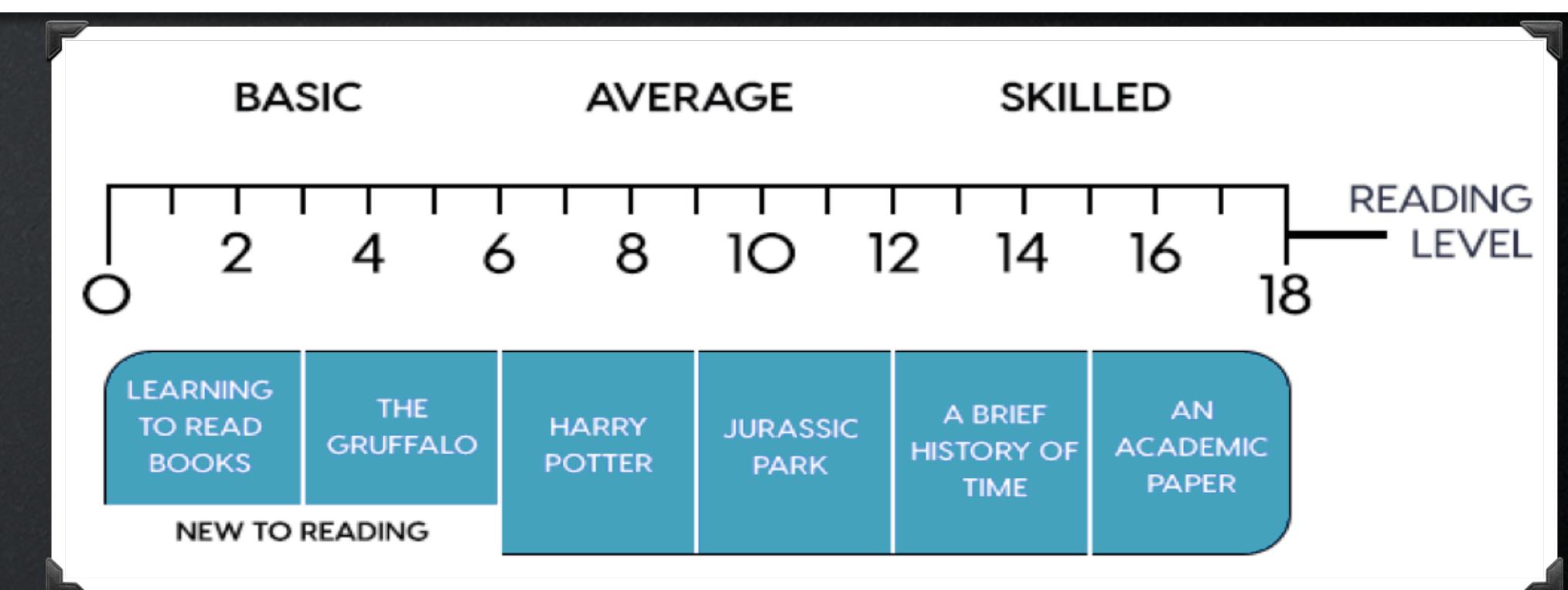
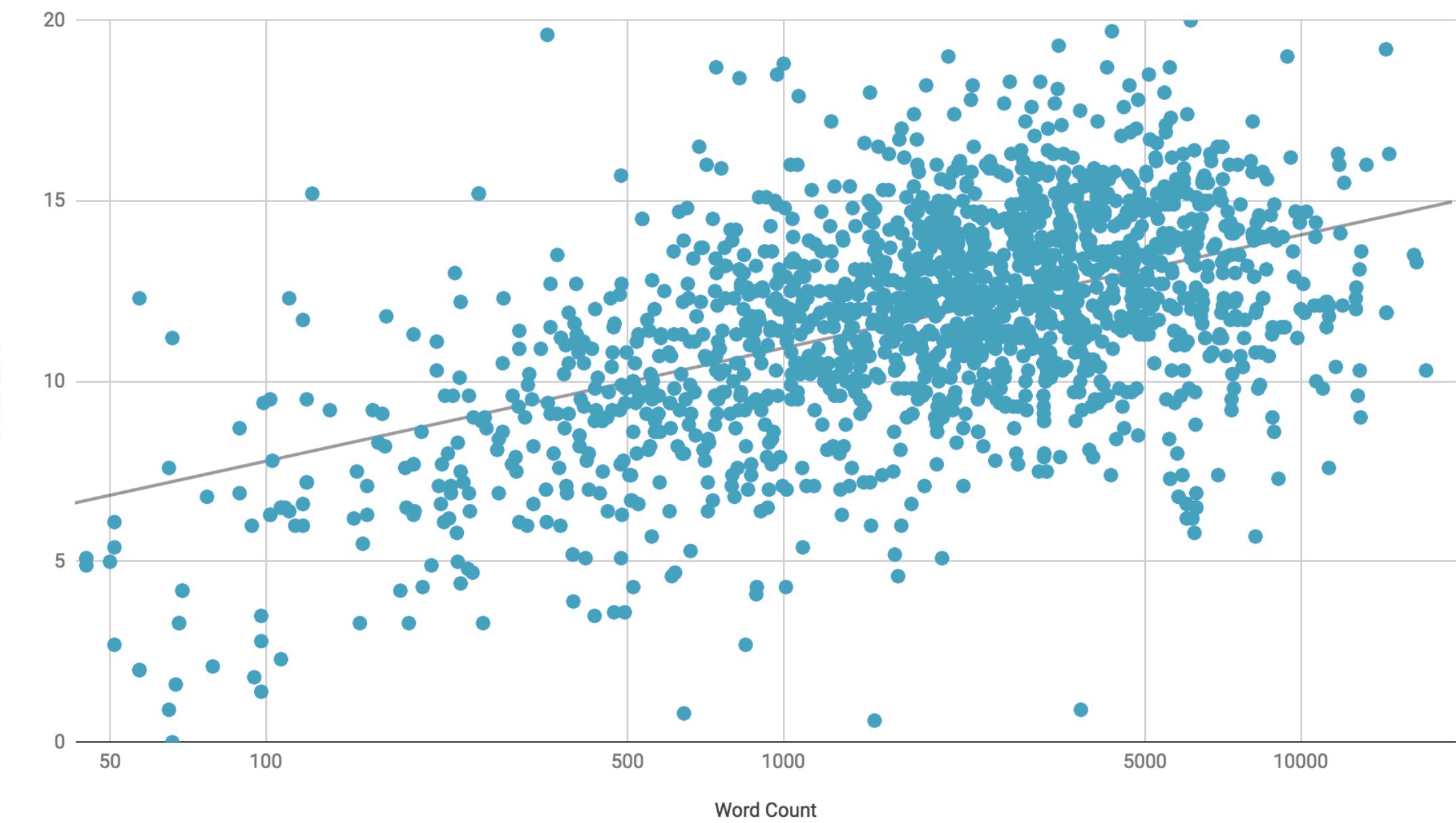


**Cause:** Increase in data privacy regulations

**Outcome:** Opaqueness in privacy policies

Grade Level vs Word Count

Grade Level vs. Word Count



Flesch-Kincaid Grade Level

([Source](#))

# Goal

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



## INFORM Module

- ★ Visualization
- ★ Annotating policies



## QUERY Module

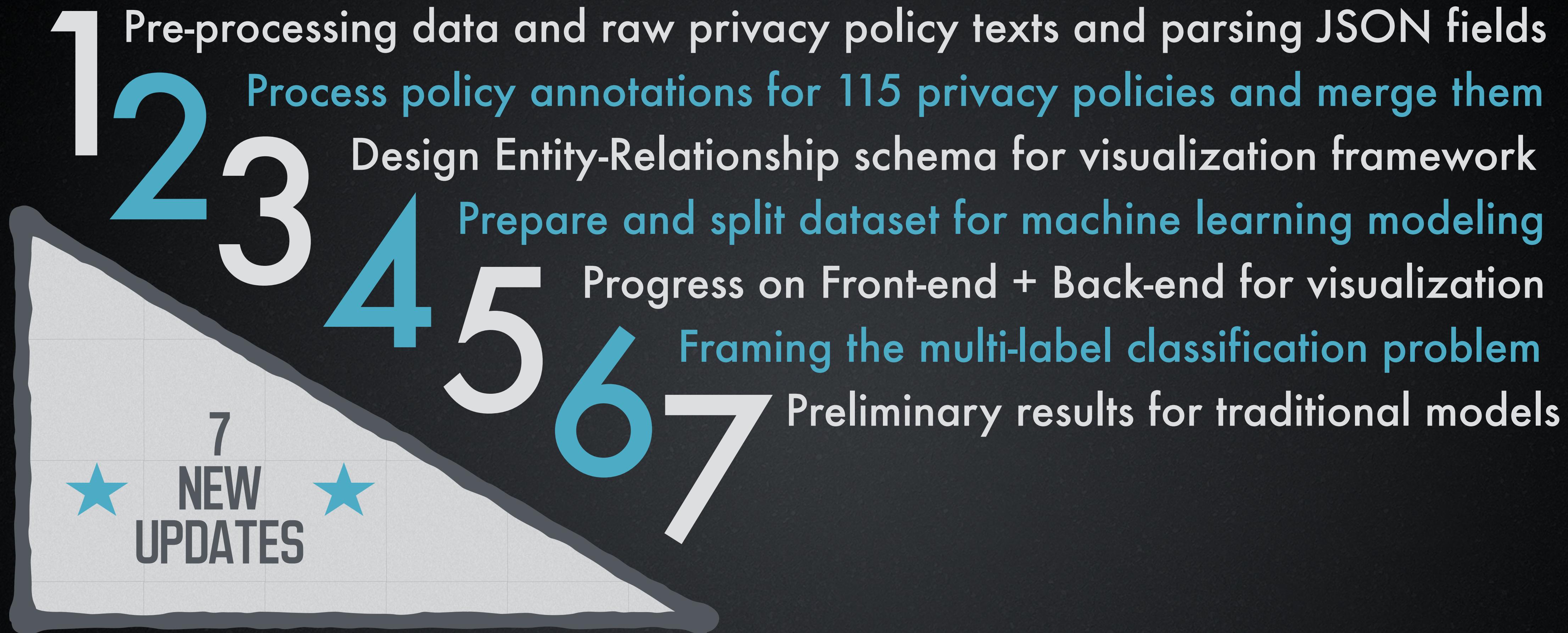
- ★ Enable users to ask policy-related

Interactive

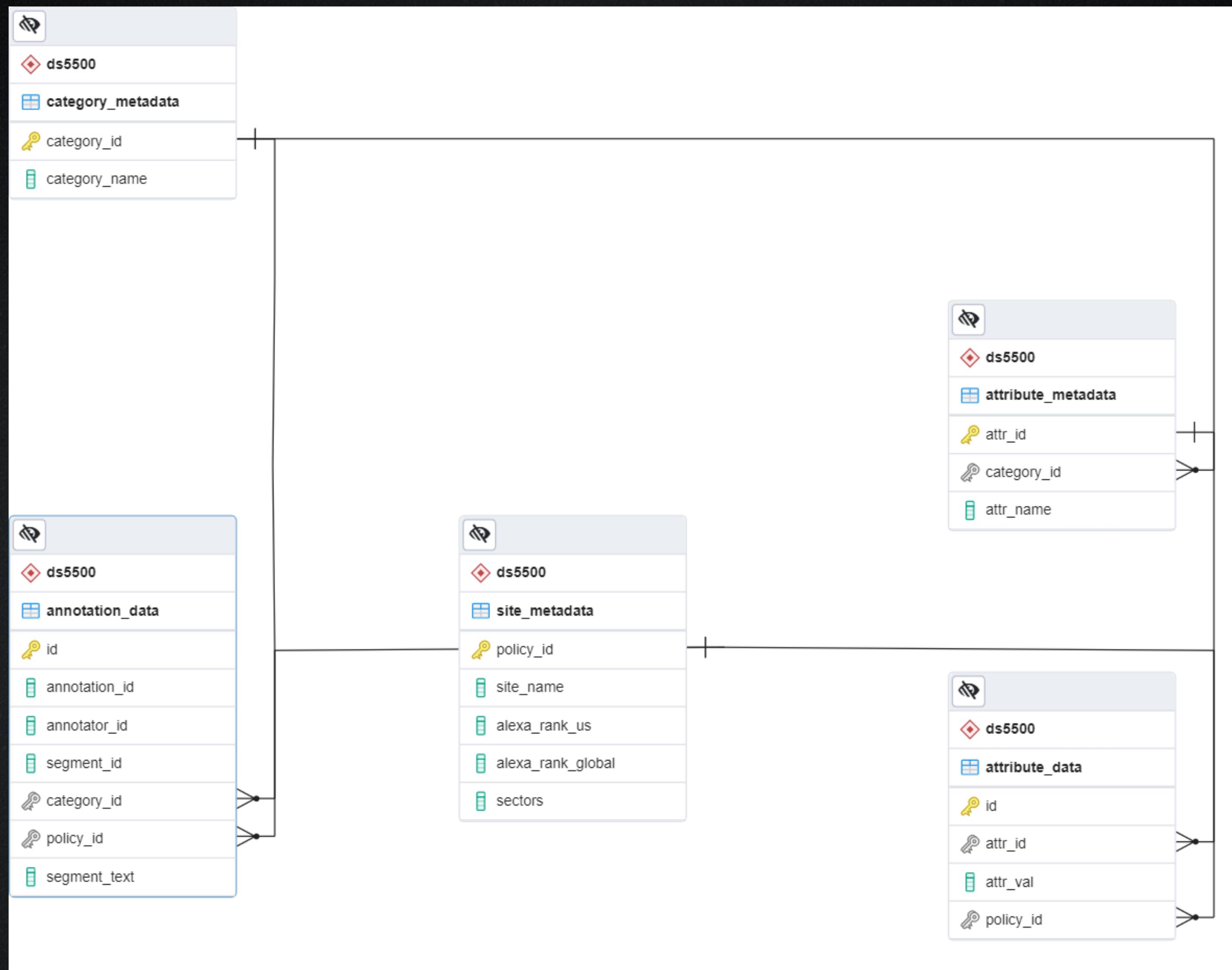
## Web Framework



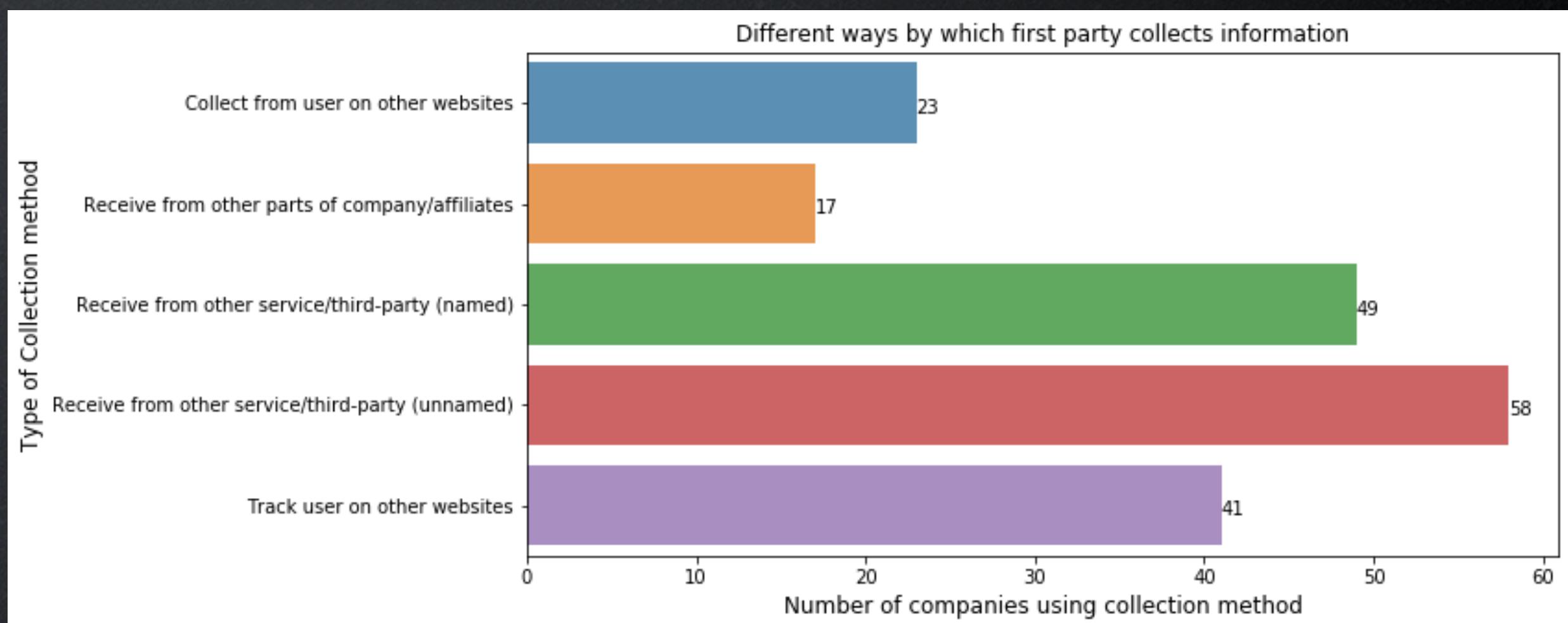
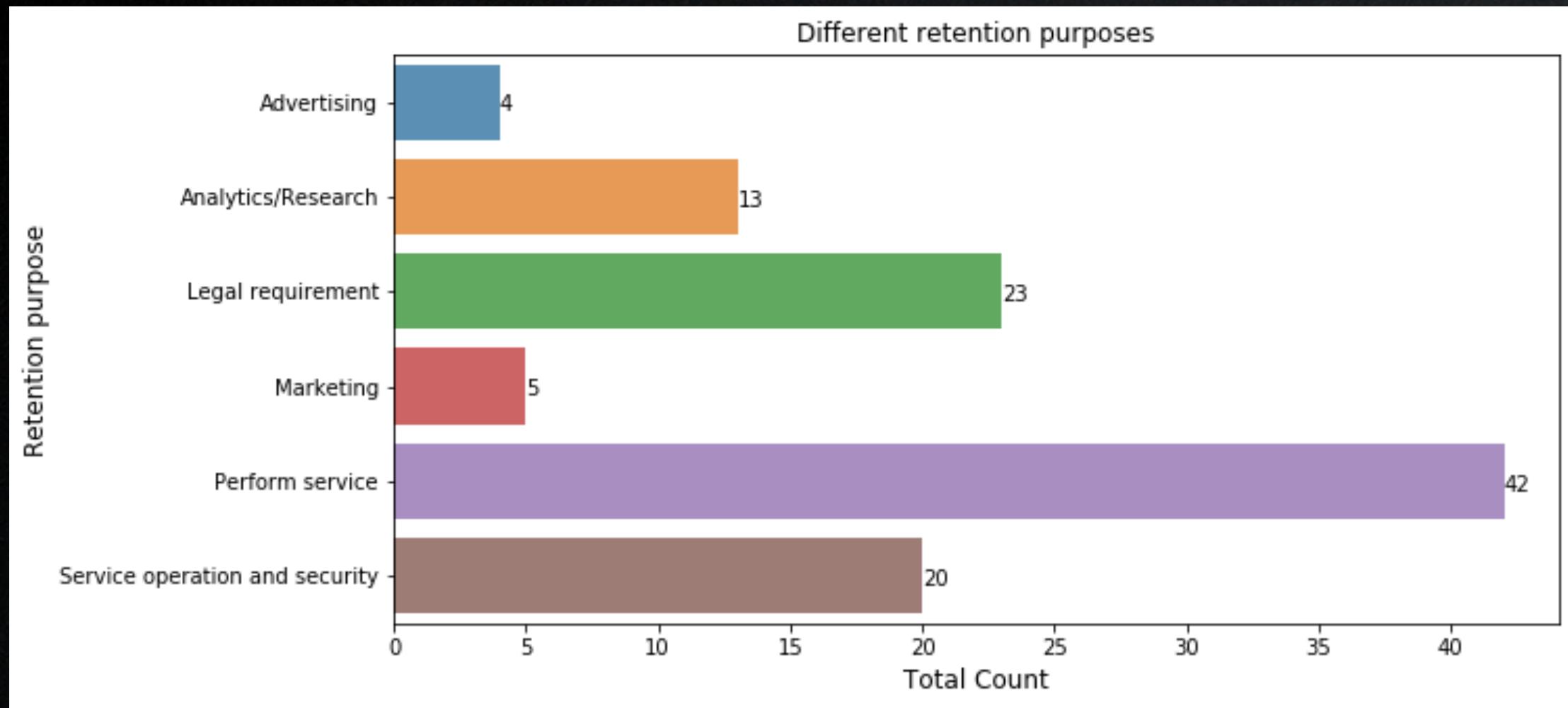
# Progress Updates



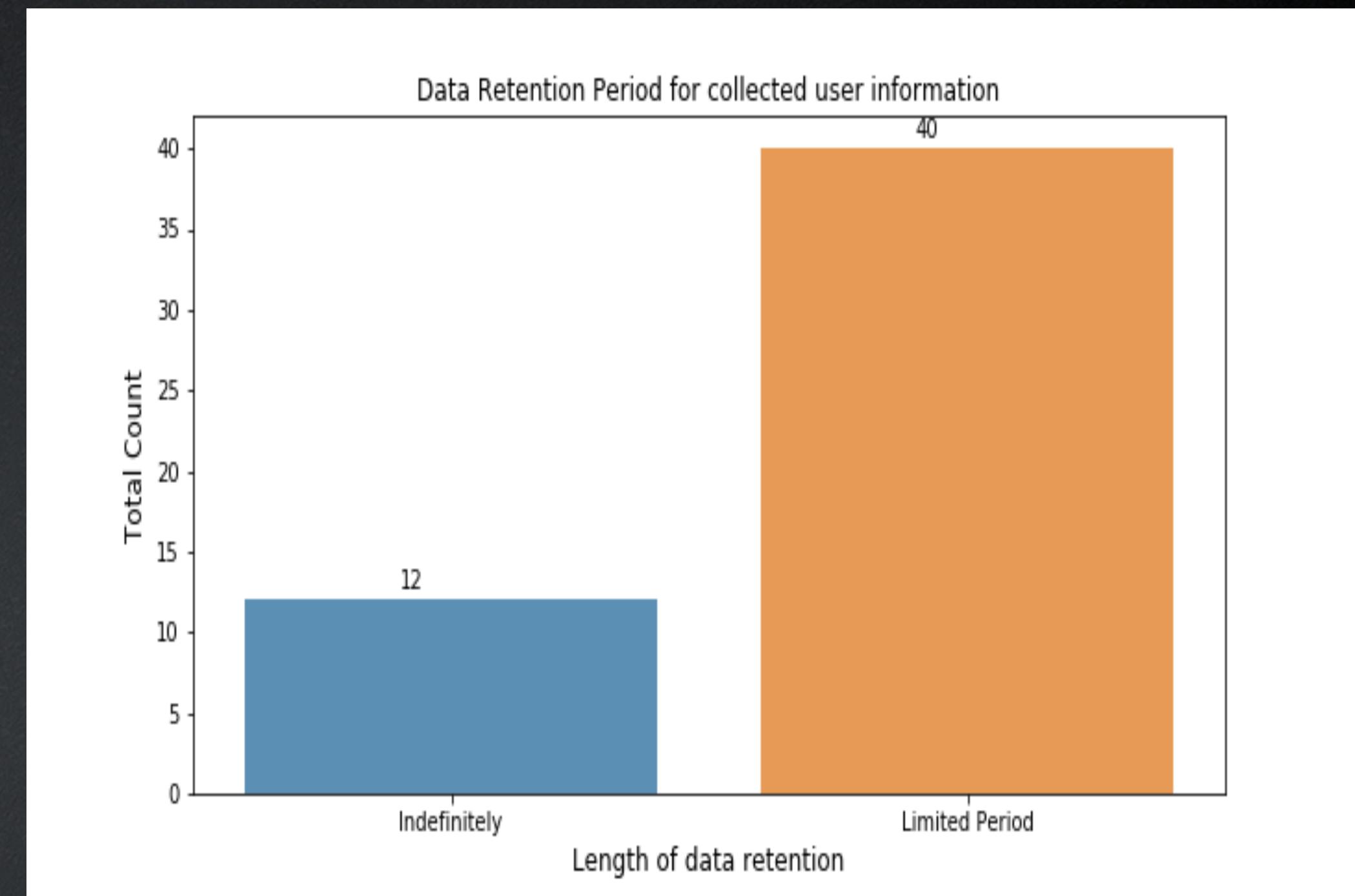
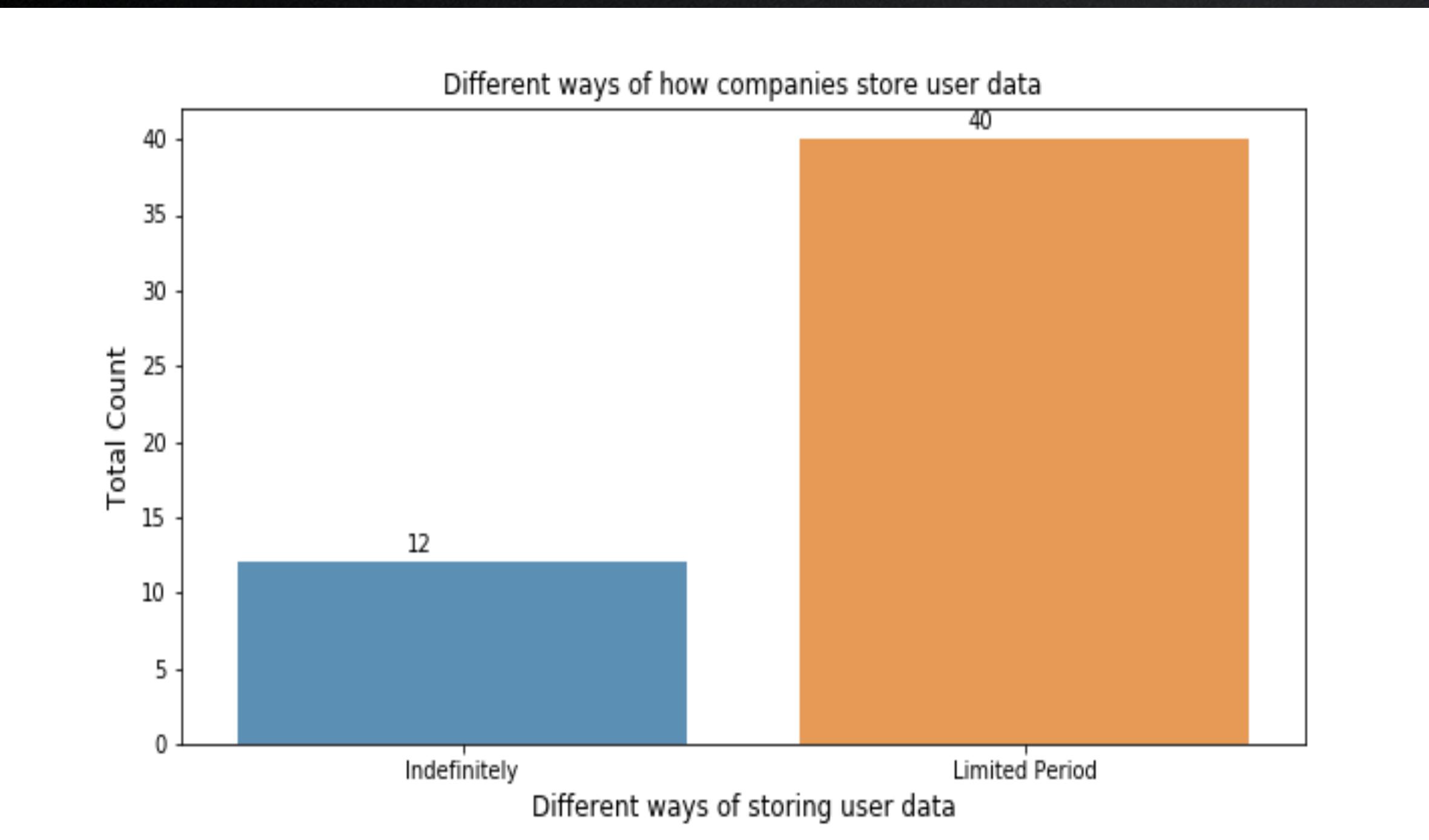
# Entity-Relationship Diagram



# Sample visualizations



# Sample visualizations (continued)



THANK  
YOU

