Predicting Compliance in Privacy Policies

Jordan Holland, Ben Kaiser, Kevin Lee, Elena Lucherini

Princeton University

CITP
CENTER FOR INFORMATION TECHNOLOGY POLICY

ABSTRACT

- Several laws that mandate what clauses commercial websites must have in their privacy policies for compliance.
- **Problem:** Sheer number of privacy policies on the Internet makes compliance enforcement through manual examination infeasible.
- **Solution:** apply machine learning methods to automatically check the compliance of privacy policies against 8 separate outcomes.

BACKGROUND AND METHOD

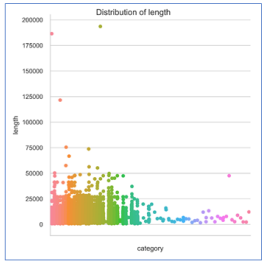
Hypothesis: Can machine learning methods automatically check the compliance of privacy policies against checklists of GDPR, CalOPPA, and CCPA requirements?

- Leverage 2 datasets:
  - 1 million privacy policies scraped from the web by CITP researchers
  - Carnegie Mellon University's OPP-115 corpus, a collection of 115 expert annotated website privacy policies
  - Train model on OPP-115 corpus and predict compliance of CITP dataset

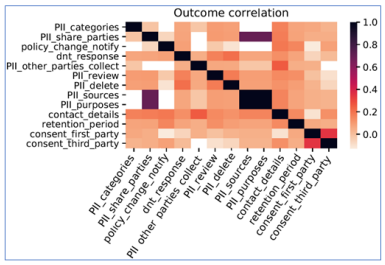
MODELING THE POLICIES

Outcome	CCPA	GDPR	CalOPPA	Comply	Don't comply	Unknown
Process to notify consumers of changes to policy	✗	✗	✓	53	13	49
Response to Do Not Track	✗	✗	✓	28	87	0
Disclose specific pieces of collected PII upon request	✓	✓	✗	41	74	0
Right to request erasure of data	✓	✓	✗	37	78	9
Contact details	✗	✓	✗	102	13	0
Retention period	✗	✓	✗	30	85	0
Subject consents to first-party processing	✗	✓	✗	25	89	1
Subject consents to third-party processing	✗	✓	✗	15	94	6

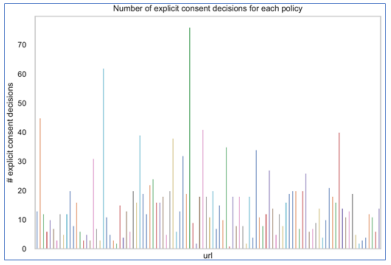
- Read and extract clauses from GDPR, CalOPPA, and CPPA that refer to concrete requirements for provisions.
- Further disambiguate the clauses, generating a set of potential outcomes
- Generate Boolean expressions over OPP-115 tags that correspond to whether the outcome was satisfied by the policy or not.
  - Select 8 outcomes (shown on the table on the left) out of original 15 with distributions most useful for prediction
- Build vocabularies for each outcome by generating bag of words for n-grams from n=1 to n=5
- Calculate 5-Fold cross validation scores for each outcome
- Predict compliance of 10,000 policies from the CITP dataset



Length distribution of grouped policies.

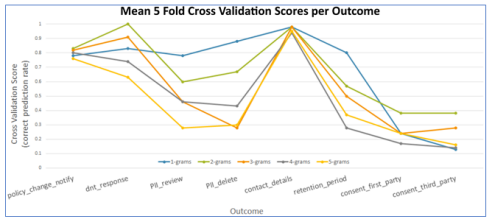


Comparison between outcomes using Pearson correlation.

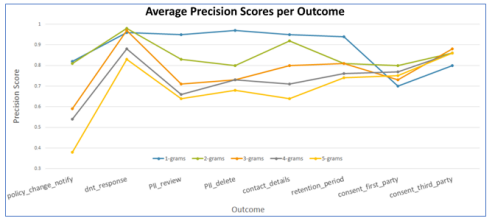


Amount of choices a user has to make when reading each policy.

RESULTS



DNT and contact detail compliance scores are over 95%, whereas third-party consent is 40%.



Average precision scores are generally higher for n=1 than for any other value of n.

Outcome	CITP policies in compliance	CITP policies not in compliance	Pct. CITP policies in compliance	Pct. OPP-115 policies in compliance
policy_change_notify	3637	6363	36%	46%
dnt_response	5241	4759	52%	24%
PII_review	9881	119	99%	36%
PII_delete	3508	6492	35%	32%
contact_details	9701	299	97%	89%
retention_period	9420	580	94%	26%
consent_first_party	4682	5318	47%	22%
consent_third_party	5763	4237	58%	13%

Compliance rates varied widely across outcomes. Generally, a larger percentage of CITP policies were compliant, as opposed to that of OPP-115 policies.

CONCLUSION AND FUTURE WORK

- Results show promise into quickly and automatically determining if a website's privacy policy is *not in compliance* by checking against some of the outcomes the classifier had higher success in predicting
- Future Work: consider word embedding on the privacy policies, which takes into the account the context of terms and may improve the performance of the classifier