Standardised Privacy Policies: A Post-mortem and Promising Developments


Reuben Binns
r@reubenbinns.com
@RDBinns
Information we collect

We collect the following categories of personal information:

- Information we collect from you: Information you provide voluntarily through our services, such as when you create an account.
- Information we collect automatically: Information collected from your use of our services, such as device information.
- Information we receive from others: Information received from other companies or organizations, such as when you link your account to another service.
- Information we infer: Information we derive from your usage data.

Information we share

We share information with the following categories:

- Google services: Information shared with Google services for product development and improvement.
- Affiliates and partners: Information shared with partners and affiliates for marketing and sales purposes.
- Government and law enforcement: Information shared with government entities for legal or regulatory compliance.
- Other uses: Information shared for other purposes as described in the Privacy Policy.

Information security

We use various security measures to protect your personal information, including encryption and access controls.

Transparency and choice

We provide transparency and choice options, such as privacy settings and user preferences.

Specific product practices

Specific product practices vary depending on the Google product.
Incentives

• Creating a standard is costly
• Adopting a standard is costly
• Extra scrutiny = loss of reputation
Network Effects

- Collective action problem
- Unilateral effort from multiple stakeholders: organisations, browsers, intermediaries, users
- Too many standards dilute efforts.
Legalese -> human -> machine

- Loss, addition or mutation of meaning
- Questionable legal standing
- Requires combination of skills from radically different domains

“License Layers” by Creative Commons, used under Creative Commons Attribution 3.0 License
“Wouldn't it be great if we had standardised templates for privacy policies?”

Platform for Privacy Preferences http://www.w3c.org/P3P (1997 – 4 years before CC!)
Khula project http://www.khulaproject.com/ (2011)
Know privacy http://www.knowprivacy.org (2009)
Netzpolitik (2007)
European privacy open space http://www.privacyos.eu (2009)
Privacy Icons (for email) http://privicons.org/ (2011)
I agree to http://www.iagreeto.org (2011)
OwnTerms http://ownterms.pbworks.com/w/page/6985495/FrontPage (2009)
Portability Policy http://portabilitypolicy.org/generator.html
Iubenda http://www.iubenda.com
PrivacyChoice http://www.privacychoice.org
Disconnect https://disconnect.me/icons
Open Notice http://www.opennotice.org
ToS;DR http://www.tosdr.org
Promising Developments

- Efficient regulatory compliance
- Digitised legal work (robo-lawyers)
- Independent / crowd-sourced ratings
- Natural language processing and machine learning
docracy
the web's only open collection of legal contracts and the best way to negotiate and sign documents online.

- find documents
- sign documents

- docracy in a nutshell
- docracy for lawyers

STANDARD LEGAL

DO-IT-YOURSELF
LEGAL DOCUMENT SOFTWARE

"That was easy! I finished my documents in just a few minutes and didn't spend money on attorney fees. Thank you!"
-Kathy P, Elyria, Ohio
Cookie Policy Generator

Having an accurate and up-to-date cookie policy, that explains exactly how you cookies on your site, is crucial to making sure you are obtaining informed consent from visitors.

On its own, it may not make your site strictly compliant with the cookie law in all parts of the EU, but if you want to avoid user intrusive pop-ups or opt-out controls, it is regarded by many as the absolute minimum any website should have.

Our Cookie Policy Generator is a simple solution to putting together an easy to

Generate a privacy policy in seconds

Start generating a privacy policy,
or scroll down to learn more ;)

You don’t want a website (yet), but you’d like to receive an update once in a while?
Leave your email!
Terms of Service
Didn’t Read

“I have read and agree to the Terms” is the biggest lie on the web. We aim to fix that.

We are a user rights initiative to rate and label website terms & privacy policies, from very good Class A to very bad Class E.

Terms of service are often too long to read, but it’s important to understand what’s in them. Your rights online depend on them. We hope that our ratings can help you get informed about your rights. Do not hesitate to click on a service below, to have more details! You can also get the ratings directly in your browser by installing our web browser add-on:
Ratings

Google [Class C]
- Google keeps your searches and other identifiable user information for an undefined period of time
- Google can use your content for all their existing and future services
- This service tracks you on other websites
- Google can share your personal information with other parties
- No promise to inform you about data requests from governments

More details

YouTube [Class D]
- Terms may be changed any time at their discretion, without notice to the user
- They can remove your content at any time and without prior notice
- The copyright license is broader than necessary
- Reduction of legal period for cause of action
- Deleted videos are not really deleted

More details

SoundCloud [Class B]
- You stay in control of your copyright
- Collected personal data used for limited purposes
- 6 weeks to review changes
- Indemnification from claims related to your content or your account
- Personal information can be disclosed in case of business transfer or insolvency

More details

GitHub [Class B]
- You don't grant any copyright license to github
- Changes can happen any time, sometimes without notice
- You shall defend and indemnify GitHub
- Your personal information is used for limited purposes
- Your account can be suspended and your data deleted any time for any reason

More details
The terms of service are very uneven or there are some important issues that need your attention.

- Terms may be changed any time at their discretion, without notice to the user
- They can remove your content at any time and without prior notice
- The copyright license is broader than necessary
- Reduction of legal period for cause of action

This service tracks and reports to Google.

You can change the terms of service any time they see fit, even without notification to the user. Your use of the service supposedly constitutes acceptance of the changes in the terms.

YouTube reserves the right to decide if a video violates the terms of service, not only if the user is a copyright infringer, but also not limited to: if the content is about pornography, obscenity, excessive length. The content may be removed and the user’s account terminated without prior notice and in the sole discretion of YouTube.

The user grants YouTube a worldwide, non-exclusive, royalty-free license that is sublicensable and transferable, and that can be used without limitation for the purpose of promoting the service, in any media format and through any media channel, provided it is in connection with YouTube, its successors’ and affiliates’.
“How fair is your contract? Find out in seconds.” - LegalSifter.com
Figure 1: Performance results for the transparency task.

<table>
<thead>
<tr>
<th>Feature</th>
<th>Average Impact</th>
<th>% of 0-Label</th>
<th>% of 1-Label</th>
<th>Feature</th>
<th>Average Impact</th>
<th>% of 0-Label</th>
<th>% of 1-Label</th>
</tr>
</thead>
<tbody>
<tr>
<td>account</td>
<td>13.09</td>
<td>100</td>
<td>100</td>
<td>information</td>
<td>-19.63</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>access</td>
<td>6.04</td>
<td>100</td>
<td>100</td>
<td>personally</td>
<td>-4.86</td>
<td>67</td>
<td>43</td>
</tr>
<tr>
<td>share</td>
<td>4.2</td>
<td>100</td>
<td>86</td>
<td>identifiable</td>
<td>-4.63</td>
<td>67</td>
<td>43</td>
</tr>
<tr>
<td>data</td>
<td>2.78</td>
<td>92</td>
<td>100</td>
<td>personally identifiable</td>
<td>-4.41</td>
<td>67</td>
<td>43</td>
</tr>
<tr>
<td>using</td>
<td>2.07</td>
<td>83</td>
<td>86</td>
<td>cookies</td>
<td>-3.58</td>
<td>92</td>
<td>71</td>
</tr>
</tbody>
</table>

Table 2: Features with the greatest-magnitude positive (left) and negative (right) impact scores.

to the classification decisions for a given test set. We measure the average impact of a feature $f$ in a train/test split of the data using the following formula:

$$\text{impact}(f) = \frac{1}{|\text{split}|} \sum_{(\text{trainset}, \text{testset}) \in \text{splits}} \text{weight}(f|\text{trainset}) \sum_{\text{doc} \in \text{testset}} \text{value}(f|\text{testdoc})$$ (1)

Despite the use of $L_2$ regularization, some frequent but not obviously informative features such as